OPTIMAL RAMP METERING OF FREEWAY CORRIDORS

A Dissertation
Presented to
The Academic Faculty

by

Bhargava Rama Chilukuri

In Partial Fulfillment
of the Requirements for the Degree
Ph.D. in the
School of Civil and Environmental Engineering

Georgia Institute of Technology
December 2015

OPTIMAL RAMP METERING OF FREEWAY CORRIDORS

Approved by:

Dr. Jorge Laval, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Angshuman Guin
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Michael Hunter
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Alen Erera
School of Industrial and Systems Engineering
Georgia Institute of Technology

Dr. John Leonard
School of Civil and Environmental Engineering
Georgia Institute of Technology

Date Approved: [November 6, 2015]

To my family

ACKNOWLEDGEMENTS

This dissertation would not have realized without my wife, Satya Bhanu Hari, who continuously supported and encouraged me during these demanding years. I can never forget the innumerable instances when she alone took care of both our children, Aditya Chilukuri and Manaswini Chilukuri, and allowed me to pursue my research. I have no doubt in my mind that she deserves special honors for being a great mother and my partner in this journey.

I am also very grateful to Dr. Jorge Laval for everything he has taught me during these years. His unmatched ability to identify means to solve complex problems, ability to blend science with art to gain valuable insights, and insatiable quest for knowledge will be a lifelong inspiration for me and one that I will strive to achieve during my professional career.

I would also like to thank Dr. Michael Hunter for teaching me how to simplify complex problems in to tangible ones and use intuitive methods to understand the underlying mechanisms. I am thankful to Dr. Angshuman Guin for several formal and informal discussions regarding research and career. His comments greatly improved the quality of this dissertation. I would like to thank Dr. Alan Erera for his detailed comments on my research contributions and dissertation that helped me address the deficiencies. I would like to thank Dr. John Leonard for his constant encouragement and guidance. Finally, I would like to thank Dr. Andrzej Swiech for patiently listening to my research problem and suggesting valuable resources.

I am also thankful to my friend Dr. Lakshmi Narasimham Peesapati for his valuable friendship that enriched my life. Our innumerable chats during lunch breaks and

weekends helped me keep moving forward. I would like to thank Hyunwoong Cho, and Aaron Greenwood for their help at several critical periods during this journey. They were two people I could always count on to get any last minute help. Thank you guys!! I would like to thank Dr. Danjue Chen, and Dr. Wonho Suh for your friendship and support. I would like to thank Carl Morris, a Ph.D candidate in ISYE, for his willingness to discuss my mathematical programming formulations and teaching me some of the common pitfalls to avoid. Finally, I would like to thank the folks at http://mathematica.stackexchange.com/ for being a great resource on Mathematica.

I would like to thank my parents, Rama Krishna Rao Chilukuri and Bala Saraswathi Bhagavatula, for everything I am today is because of their unconditional love.

I also want to thank my brother Raghavendra Chilukuri and sister-in-law Sowjanya Pisipaty for their constant encouragement. I would like to thank my parent-in-laws, Dr. Rambabu Hari and Indira Hari and my brother in-law, Murali Krishna Hari for everything they have done to help me smoothly navigate this arduous journey.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS IV

LIST OF TABLES IX

LIST OF FIGURES X

LIST OF SYMBOLS AND ABBREVIATIONS XIII

SUMMARY XIV

CHAPTER 1 INTRODUCTION 1

1.1. Research Objectives 2

1.2. Overview 3

CHAPTER 2 LITERATURE REVIEW 5

2.1. Isolated Metering 6

2.2. Coordinated Metering 8

2.2.1. Optimal Control Methods 8

2.2.2. Rule-Based Methods 11

CHAPTER 3 BACKGROUND 14

3.1. Fundamental Diagram 14

3.2. Merge Model 15

3.3. Traffic Flow Models 17

3.2.1. First-Order Models 17

3.2.1.1. Godunov Scheme 18

3.2.1.2. Variational Theory 21

3.2.2. Second-Order Models 23

3.4. Optimal Control Theory 24

CHAPTER 4 ANALYTICAL FORMULATIONS 28

4.1. Discrete Ramps with Unlimited Queue 28

4.2. Continuum Ramp Formulation under Queue Constraint 32

CHAPTER 5 DISCRETE TIME-SPACE FORMULATIONS 35

5.1. Exact Godunov Scheme 37

5.2. Simulation-Based Optimization 40

5.3. LP Formulation 43

CHAPTER 6 ANALYSIS WITH LP FORMULATION 47

6.1. Monotonic demand 48

6.2. Dimensionless Formulation 58

6.3. Merge Model Impacts 61

CHAPTER 7 HYBRID ISOLATED CONTROLLER 63

7.1. Philosophy 63

7.2. Application 65

CHAPTER 8 DISCUSSION AND EXTENSIONS 69

8.1. Contributions 70

8.2. Extensions 71

APPENDIX A. OPTIMALITY CONDITIONS WITH ALTERNATIVE FORMULATION 73

REFERENCES 74

LIST OF TABLES

TABLE 1: PARAMETERS FOR NUMERICAL EXPERIMENT 47

LIST OF FIGURES

FIGURE 1: FREEWAY-RAMP CONFIGURATION AND PARAMETERS 7

FIGURE 2: FUNDAMENTAL DIAGRAM A) GREENSHIELDS B) TRIANGULAR 15

FIGURE 3: NEWELL-DAGANZO MERGE MODEL 17

FIGURE 4: DEMAND AND SUPPLY FUNCTIONS 20

FIGURE 5: SAMPLE DEMAND FUNCTIONS A) HADIUZZAMAN ET.AL. [92] B) LEBACQUE [75] 20

FIGURE 6: VISUAL REPRESENTATION OF VARIATIONAL THEORY SOLUTION TO THE KW MODEL 23

FIGURE 7: CORRIDOR ARRIVAL AND DEPARTURE N-CURVES 28

FIGURE 8: ARRIVAL AND DEPARTURE CURVES 29

FIGURE 9: TIME-SPACE GRID FOR DISCRETE FORMULATION 36

FIGURE 10: FLOW CALCULATION USING EXACT GODUNOV METHOD 39

FIGURE 11: SHOCK PROPAGATION WITH CLASSICAL GODUNOV AND EXACT GODUNOV METHODS 39

FIGURE 12: SIMULATION-BASED OPTIMIZATION FRAMEWORK 40

FIGURE 13: TIME-SPACE METERING RATE FOR a =2000 VEH/HR-KM AND b =0.1/KM 48

FIGURE 14: RELATIONSHIP BETWEEN OPTIMAL CONTROL STRATEGY AND STATE FOR a =2000 VEH/HR-KM AND b =0.1/KM 49

FIGURE 15: OPTIMAL METERING AND RAMP QUEUES IN ZONE A 50

FIGURE 16: OPTIMAL METERING AND RAMP QUEUES IN ZONE B 51

FIGURE 17: SHOCK SPEEDS ON THE RAMPS DURING OPTIMAL CONTROL... 52

FIGURE 18: OPTIMAL METERING AND RAMP QUEUES IN ZONE C 53

FIGURE 19: OPTIMAL METERING AND RAMP QUEUES IN ZONE C – SCENARIO 1..... 54

FIGURE 20: OPTIMAL METERING AND RAMP QUEUES IN ZONE C – SCENARIO 2..... 55

FIGURE 21: OPTIMAL METERING AND RAMP QUEUES IN ZONE D 56

FIGURE 23: RELATIONSHIP BETWEEN OPTIMAL CONTROL STRATEGY AND STATE FOR INCREASING a AND DECREASING b..... 57

FIGURE 24: $x0$ IN RESCALED DIMENSIONLESS VARIABLES 60

FIGURE 25: $s1$ IN RESCALED VARIABLES..... 60

FIGURE 26: REGION SATISFYING RAMP PRIORITY ON THE MERGE MODEL DIAGRAM A) PARAMETERS B) MERGE MODEL..... 62

FIGURE 27: BLOCK DIAGRAM OF THE HYBRID ISOLATED CONTROL..... 64

FIGURE 28: CUMULATIVE ARRIVALS AND DEPARTURES AT THE RAMP METER; NO METERING, ALINEA, AND HYBRID CONTROL..... 66

FIGURE 29: TIME-SPACE PLOT OF THE DISCRETE RAMPS. A) NO METERING B) ALINEA C) HYBRID CONTROL 66

FIGURE 30: TIME-SPACE DENSITY PLOT A) ALINEA B) HYBRID CONTROL .. 67

LIST OF SYMBOLS AND ABBREVIATIONS

FD	Fundamental Diagram
KW model	Kinematic Wave Model
LWR model	Lighthill Witham Richards model
CTM	Cell Transmission Model
EG scheme	Exact Godunov scheme
NLP	Nonlinear Programming
MILP	Mixed Integer Linear Programming
LP	Linear Programming

SUMMARY

Ramp meters have been used for congestion management on freeways since the 1960s to maximize freeway capacity by controlling on-ramp flows. Traditionally, the focus has been to develop rule-based algorithms and optimal control case studies. This led to a host of algorithms and methods which cannot be proven to provide an optimal control and the case studies does not provide a systematic understanding of the characteristics of optimal control and its influence on traffic dynamics. Moreover, optimal is not easy to achieve in practice due to the limited storage on the on-ramps.

Towards this end, this dissertation systematically studies the optimality conditions for the case of unlimited storage and spatiotemporal evolution of control and its corresponding traffic dynamics on freeway and ramps under queue constraint, carefully taking the traffic dynamics into account.

A Kinematic Wave model of the freeway-ramps system is optimized for minimal total delay. The optimality conditions for the case of unlimited ramp storage are studied using Moskowitz functions that provide several interesting insights for different scenarios, including the case of limited storage. This dissertation shows that the current problem posed as a nonlinear coupled PDE system with a nonlinear merge model cannot be solved analytically. This study also shows that the discrete-time nonlinear formulation solved with simulation-based optimization does not converge in reasonable time. To overcome this, the problem is reposed as a LP formulation that includes capacity drop. For discrete formulation, this study develops an error-free solution to the KW model with a source term that enhanced the quality of the numerical solution. This study identifies four distinct regions in the state surface with distinct metering patterns. Explicit modeling

of ramps enabled correlating the initialization and termination times of the metering patterns with the evolution of traffic dynamics on the freeways and ramps. Using these results, this dissertation presents a hybrid isolated ramp metering algorithm that outperforms existing methods.

CHAPTER 1 INTRODUCTION

Ramp meters have long been implemented as freeway congestion management tool to maximize freeway capacity by controlling inflows at on-ramps. Primary benefits of ramp meters are two-fold: prevent capacity drop and avoid exit blockage [1]. They also provide additional benefits such as diversion to alternate routes, enhanced safety, and reduced emissions.

Generally, on-ramps are controlled independently [2]–[6] or in coordinated groups [7]–[13] (see [14], [15] for summaries). Isolated ramp meters operate based on the traffic conditions in its vicinity and the coordinated control manages multiple on-ramps based on the traffic conditions within its control region. The objectives commonly used for the ramp control problems are maximizing freeway flow and minimizing total system delay (both freeway delay and ramp delay).

Assuming that there is enough demand in the system to create freeway congestion, a feasible solution to the ramp control problem may not be difficult to find since any metering of the on-ramp flows is bound to alleviate freeway congestion, which in turn will increase freeway flows and reduce total delay. However, obtaining an optimal solution that minimizes combined delay of both components of the system (freeway and ramps) is difficult due to a large number of feasible metering combinations. Additionally, limited storage on the on-ramps increases the complexity of the optimal control problem since restrictive metering often results in queue spillback to the arterials, which is often unacceptable. Thus, incorporating ramp storage into control problem mean that the optimal control should not only minimize total delay, but also satisfy the queue constraint.

Traditionally, the focus of the isolated ramp metering research has been to develop algorithms that aim to maintain a target state on the freeway [6], [16], [17], which will lead to suboptimal performance in the absence of knowledge of optimality conditions.

The focus of the coordinated control research has been to develop rule-based algorithms [5], [9], [12], [18]–[22] and optimal control case studies [23]–[29]. The former lead to a host of heuristic methods that are generally complex, computationally intensive, expensive field infrastructure dependent, and sensitive to data detector failures and still cannot be proven to provide optimal control. Also, the optimal control case studies only provide limited insight into generalized spatiotemporal aspects of optimal control and its corresponding traffic dynamics on freeway and ramps. This is not surprising since the optimal control problem involves multiple variables such as spatiotemporal distribution of on-ramp and off-ramp demands, time-varying mainline flow, freeway and ramp geometry, nonlinear nature of flow-density relationship capacity drop, nonlinear merging behavior, capacity drop, etc.

To fill this gap, this research systematically studies the optimal control problem for minimal total delay. The optimality conditions for the case of unlimited storage and spatiotemporal evolution of control and its corresponding traffic dynamics on freeway and ramps under queue constraint, are studied carefully taking the traffic dynamics into account. Towards this end, the next section will present the scope of the problem studied in this research.

1.1. Research Objectives

This research aims to study the optimal control problem for both unlimited and limited storage cases. We show that for the case of unlimited storage on the ramps, the

optimality conditions can be obtained when the system is modeled using Moskowitz functions. For the case of queue constraint, system modeling and rigorous optimization are required to obtain optimal solutions. Towards this end, this research aims to address the following:

- Study the optimality conditions for the case of unlimited storage on the on-ramps.
- Investigate the application of the simulation-based optimization framework for solving the optimal control problem;
- Unveil the spatiotemporal characteristics of optimal control and identify their relationship with the traffic dynamics on freeway and ramps for the case of constrained on-ramps queues; and
- Use the insights to develop an efficient ramp metering method.

1.2. Overview

The remainder of the dissertation is organized as follows:

The next chapter presents a literature review of the state-of-the-art and state-of-the-practice isolated and coordinated ramp metering methods. The chapter will also describe the limited insights gained from the optimal ramp metering case studies.

Chapter 3 presents a brief background on a) first-order and second-order traffic flow models and their numerical solutions, and b) optimal control theory methods for continuous time and discrete time-space formulations.

Chapter 4 presents the analytical formulations using Moskowitz functions for the case of discrete ramps, which provides an insight into the optimality conditions for the case of unlimited ramp queues. The later part of the chapter presents the continuous ramp

formulation using the Kinematic Wave (KW) model and shows why such a formulation cannot be solved analytically.

Chapter 5 present three discrete time-space formulations (nonlinear, mixed integer, and linear programming) and describes their limitations for solving the optimal control problem. The chapter also presents a superior numerical solution scheme for the KW model with source term, which is used in the simulation-based optimization framework. Finally, the LP formulation used for the optimal control problem is presented.

Chapter 6 utilizes the LP formulation from the previous chapter to investigate the spatiotemporal optimal control trajectory and the state trajectory on the freeway and ramps to identify patterns and relationships. Analytical solutions for initiation and termination of the metering patterns, governed by the shockwave evolution are presented. The chapter also presents the dimensionless formulations and merge model impacts to demonstrate some insights and limitations of the results.

Chapter 7 presents an isolated hybrid feedback-feedforward control, which combines the strengths of both types of controllers and outperform existing methods.

Chapter 8 presents discussion, contribution, limitations and extensions of the current work.

CHAPTER 2 LITERATURE REVIEW

Ramp meters break up the vehicle platoons entering the freeway to enable smooth merging of the on-ramp flows to minimize total system delay. They also control excessive inflows into the freeway to avoid freeway queues from blocking the off-ramps.

Ramp metering control began with independent metering at individual on-ramps to control isolated bottleneck congestion that arise due to the merging flows. Isolated metering is simple and economical, but cannot handle severe and/or widespread congestion. Coordinated control is used when there are multiple bottlenecks within a region or congestion extends to a larger region. Coordinated control manages metering rates at multiple ramps to efficiently store vehicles, but is complicated to efficiently implement in practice.

As mentioned in CHAPTER 1, determining optimal metering rates that minimize delay is difficult. Towards this end, researchers developed some general theories to provide guidance for the practitioners. Wattleworth [30] suggested that local meters should aim to maintain the freeway at critical density to minimize delay. Pretty [31] suggested that spatial distribution of control should be ordered such that the fraction of their flow destined for the bottleneck is decreasing. However, a systematic understanding of optimal control characteristics is missing.

The rest of the chapter presents a brief overview of the state-of-the-art and state-of-the-practice of the isolated and coordinated ramp metering methods.

2.1. Isolated Metering

Isolated metering strategies are classified into 4 categories- based on the underlying methods: linear programming [3], [4], [32], control theory [5], [6], [16], [21], neural networks [33], and fuzzy-logic [18], [34]. Of these, methods based on linear programming and control theory are popular and field implemented extensively.

Fixed-time plans began with the work of Wattleworth and Berry [32] who maximized metering rates under free-flow conditions using linear programming. However, the limitation with their formulation is that it does not consider the nonlinearity of the traffic flow. Other researchers [3], [4], [35]–[38] built on this methodology and extended their formulation for user-optimal flow, ramp queues, congested state, etc. The limitation with these efforts is that their plans are developed based on historic data and does not respond to traffic conditions in real-time.

Masher et. al. [16] developed the first traffic responsive control for isolated metering: demand-capacity and percentage occupancy methods. These methods are based on the idea that the metering rates are set to aggregate the merge flows to capacity flow until the merge density exceeds critical density, as explained below.

Consider a freeway-ramp section and the traffic parameters as shown in Figure 1, where $q(.)$ and $r(.)$ represents the average flow and $o(.)$ average occupancy. The demand-capacity strategy states that the metering rate at time t is:

$$r(t) = \begin{cases} Q - q_{in}(t-1), & O_{out} < O_{cr} \\ q_{min}, & otherwise \end{cases} \quad (1)$$

where Q and O_{cr} are the desired flow and occupancy (which are typically set to capacity

and critical occupancy) and q_{min} is the minimum metering rate.

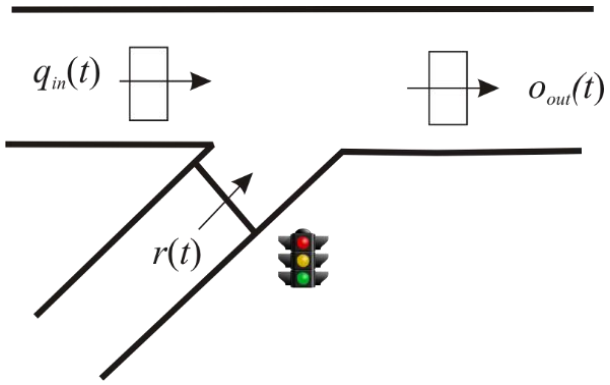

Figure 1: Freeway-ramp configuration and parameters

The philosophy of the percentage-occupancy strategy is similar to the demand-capacity strategy and only differs in its implementation. Instead of measuring upstream flow, $q_{in}()$, the percentage-occupancy strategy uses upstream occupancy to estimate the flow. Moreover, the upstream occupancy is used for the switchover decision. Therefore, the percentage-occupancy strategy only needs one freeway detector (upstream of the ramp location) to determine metering rate.

These methods were field implemented extensively and found to perform better than the fixed-time plans [16]. The strengths of these methods is that they quickly respond to any under-capacity conditions and are known to be sensitive to unmeasured disturbances, i.e. moving bottlenecks, shockwave. Their drawback is that the control changes significantly and drastically only after the congestion sets in. Also, minimum rate may sometimes lead to unnecessary under-loading of the freeway.

To overcome these drawbacks, Papageorgiou et. al. [5], [17] developed ALINEA, which belongs to the class of integral-feedback controller that corrects for the cumulative

error. ALINEA aims to maintain critical occupancy at the merge location as follows:

$$r(t) = r(t-1) + K_R(O_{cr} - O_{out}(t)) \quad (2)$$

where K_R is the regulating parameter, which is in the range of 60-70 veh/hr. ALINEA was field implemented extensively and found to perform better than the fixed-time plans [17], [39]. ALINEA feedback law is found to be simple, flexible, robust, and provides smooth transitions during congestion build up and dissipation. However, the reactive nature of ALINEA results in lag in responding to large disturbances and takes several iterations to restore back to the desired state.

To overcome implementation issues, several variations of ALINEA such as AD-ALINEA (adaptive strategy to dynamically calculate critical occupancy), AU-ALINEA (upstream-measurement-based version of the AD-ALINEA), FL-ALINEA (flow-based ALINEA), UP-ALINEA (upstream-occupancy-based), UF-ALINEA (upstream-flow-based), and X-ALINEA/Q (combination of any of the preceding strategies with queue control) were developed [6], [40].

2.2. Coordinated Metering

Unlike isolated control that manages each ramp independently, coordinated control manages a group of on-ramps based on traffic conditions within its control region. Coordinated metering methods can be broadly divided into two categories: optimal control methods, and rule-based control methods.

2.2.1. Optimal Control Methods

Optimal control methods can be divided into the ones that use optimization methods and that use automatic control theory methods.

Optimization-based coordinated metering began with Wattleworth and Berry [32] and Wattleworth [30] who extended the linear programming-based formulation for the isolated control to coordinated control. They maximized the metering rates under steady-state conditions for the corridor to determine fixed-time metering plan for a group of on-ramps. Yuan and Kreer [3] build on the Wattleworth's methodology to maximize the flow and balance the ramp queues. Papageorgiou [38] and Banos and Papageorgiou [37] developed optimal control by relaxing the steady-state condition. Lovell and Daganzo [41] used time-dependent origin-destination flows for free-flow conditions to develop a heuristic access control algorithm for special cases. The limitation of these methods is that they do not capture the nonlinearity of the traffic flow.

Recently, Gomes and Horowitz [42] proposed Asymmetric Cell Transmission Model to derive optimal control for both free-flow and congested conditions. They circumvent the non-linearity in the traffic flow by posing the flows as inequality and their objective function ensured that the flows take the upper bound. This study adopts their methodology, but extends/differs in some ways to understand the relationship between the optimal control trajectory and state evolution of freeway and ramps.

The control theory-based methods began with Isaksen and Payne [43] who formulated a linear regulator problem with quadratic cost to develop a feedback control rule. Goldstein and Kumar [44] developed a multilevel decentralized control scheme based on linear quadratic regulator that outperformed the centralized control. Chen et. al. [45]

developed a hierarchical dynamic control by posing local and area-wide metering control as a nonlinear control problem. Kotsialos et.al. [25] presented a generic approach to solving the ramp control problem by formulating it as a constrained discrete-time nonlinear optimal control problem, which was used in several studies based on second-order models [26], [46]–[48].

Only limited optimal coordinated control methods were field implemented. Diakaki and Papageorgiou [49] and Papageorgiou et.al. [50], [51] developed a multivariate control, called METALINE, to efficiently manage multiple ramps. METALINE is a vectorized extension of ALINEA as shown below:

$$\boldsymbol{r}(t) = \boldsymbol{r}(t-1) + \boldsymbol{K_1}(\boldsymbol{o}(t) - \boldsymbol{o}(t-1)) + \boldsymbol{K_2}(\widehat{\boldsymbol{O}} - \boldsymbol{O}(t)) \qquad (3)$$

where $\boldsymbol{r}(k)$ indicates a set of controlled on-ramps $\boldsymbol{r} = [r_1, r_2, \dots r_n]^T$, $\boldsymbol{o}(t)$ indicates a set of state measurements $\boldsymbol{o} = [o_1, o_2, \dots o_m]^T$, and $\boldsymbol{O}(t)$ indicates a subset of state measurements $\boldsymbol{o}$ for which target occupancies are available $\widehat{\boldsymbol{O}} = [\hat{O}_1, \hat{O}_2, \dots \hat{O}_n]^T$. Finally, $\boldsymbol{K_1}$and $\boldsymbol{K_2}$ are the calibrated gain matrices. Field applications in Paris [50], [52] and Amsterdam [49] indicated that METALINE is simple and robust, but its performance is highly sensitive to the static gain matrices that are calibrated using historic data.

The limitation of the above efforts is that their focus is on the problem formulation and application. To overcome this, Zhang et.al. [8], [53] used KW model formulation to derive some general analytical results for ramp metering with and without alternate routes. Zhang [8] noted that ramp metering is not beneficial when traffic is uniformly congested, or uniformly uncongested, but effective when traffic conditions are transient. They also argue that ramp metering may not reduce total delay unless drivers use alternate routes.

In addition to the optimal control for coordinated ramp metering, some researchers developed integrated control of ramp metering with traffic signals [7], variable speed limits [27], [54], route guidance [9], [26], and other [28], [46], [48], [55], whose focus was predominantly on developing synergic control.

Note that the control trajectory obtained using optimal control methods is open-loop and one needs to use Model Predictive Control-like frameworks to develop real-time optimal traffic responsive systems [9], [26]. Thus, the optimal control methods need complex numerical solution algorithms that are computation intensive. Therefore, the state-of-the-practice has been to use the rule based metering that are described in the next subsection.

2.2.2. Rule-Based Methods

Rule-based algorithms are popular and field implemented extensively. They use ad hoc methods to calculate ramp metering rates for a group of on-ramps in real-time. Some of the popular rule-based algorithms include Minnesota Zone algorithm [10], SWARM [11], HERO algorithm [12], Seattle Bottleneck algorithm [13], Linked-ramp metering algorithm [56], and Denver Helper algorithm [57].

The Minnesota Zone algorithm divides the freeway into zones with their upstream boundary in free-flow and downstream boundary a bottleneck. The algorithm uses vehicle conservation to ensure that the total entering volume is at most equal to the total exiting volume. The metering rates at the ramps are calculated based on the existing capacity of the bottleneck, predetermined importance of the ramp, local traffic conditions, etc. By comparing the local metering rate and the corresponding downstream occupancy, a more

restrictive metering rate is applied. The Stratified Zone algorithm, an improved version of the Zone algorithm, balances the freeway efficiency and ramp delay to maximize freeway flow. Field evaluation showed that the traffic volumes and safety increased and the travel time and emission are decreased by 9%-14% [58].

Similar to the Zone algorithm, SWARM divides the corridor into segments and all the ramps within a segment are controlled as a group. The control operates on global and local level, with the former doing the forecasting and apportioning and the latter providing local responsive metering. The more restrictive of the two metering rates is implemented at each ramp. Empirical studies by Pham [59] indicated that SWARM increased mainline speed by 11%, cut down travel time by 14%, and delay by 17%.

In the HERO algorithm, each ramp is outfitted with an ALINEA algorithm. When the ramp queue exceeds a threshold it becomes a "master" control and manages some upstream ramps to reduce the queues at the Master ramp. The aim of this algorithm is to efficiently use all the space available on the network, but does not optimize for the freeway-ramp system. Field evaluation showed 8.4% increase in average flow and a 58.6% increase in average speed [12].

The Seattle Bottleneck algorithm aims to maintain the flows at the bottleneck locations under saturation level. The freeway is divided into sections, where local metering rates are determined by the demand-capacity type of strategy and the global metering rates are computed based on the bottleneck density and required volume reduction. The volume reduction is distributed to the associated upstream on-ramps according to pre-determined weighting factors. Then, the more restrictive one will be selected for the field

implementation. Global metering rates are applied only when the downstream occupancy is above a threshold or the number of entering vehicles through freeway and on-ramps is greater than the number of exiting vehicles via freeway and off-ramps.

The Linked-Ramp Metering system uses demand-capacity approach or determining the ramp metering rate. Occupancy measurements are used to identify congestion and correct the erroneous flow measurements during congestion. The coordination achieved heuristically by restricting upstream ramps when downstream ramps fill up.

The Helper Algorithm is a hierarchical coordinated control method where local metering rates are calculated using demand-capacity type methods. The coordination is achieved heuristically based on the traffic conditions and the queue lengths on the on-ramps. Similar to other rule-based methods, this algorithm also restricts upstream ramp flows when the downstream ramps grow longer.

The drawback with these methods is that they all generally employ ad hoc feedforward control to achieve a target flow. As Newman [60] argues, the observed benefits of rule-based algorithms may be an artifact of the decision to have a metering system and not the quality of the algorithm, which reinforces the need for a better understanding of the optimality conditions.

This chapter presented a literature review of the ramp control methods, which indicated that there is limited understanding of the qualitative aspects of optimal control. Towards addressing this, the next chapter will present background information on the traffic flow models used in this study.

CHAPTER 3 BACKGROUND

This chapter presents the existing background material upon which this dissertation is built. It is divided into four sections. The first section presents a brief overview of the origin and evolution of the fundamental diagram. The second section describes the Newell-Daganzo merge model [61]. The third section presents the first-order and second-order traffic flow models and their numerical solutions. The last section presents the optimal control theory methods for continuous time formulation.

3.1. Fundamental Diagram

Traffic flow is generally represented by: a) density, in vehicles/unit distance - $k(t,x)$, b) flow, in vehicles/unit time - $q(t,x)$, and c) velocity, in distance/time - $v(t,x)$, related by the following fundamental relationship:

$$q(.) = k(.)\, v(.) \qquad (4)$$

In the traffic theory, the relationship between any two traffic variables (of the three mentioned above) is called a fundamental diagram (FD). Based on empirical observations, Greenshields [62] proposed the first FD, which gives a linear relationship between speed and density. Inserting Greenshields' speed-density relationship into *(4)* yields a unimodal function $q = F(k)$, with $F(0) = F(\kappa) = 0$, where κ is the jam density (see Figure 2a) and maxima at k_c.

Over the years, researchers have proposed several other FDs (see [63] for a summary). Recent empirical studies [64]–[68] showed that the acceleration and deceleration waves travel at nearly constant speeds indicating the shape of the FD to be triangular (see Figure 2b). This FD represents the generally accepted two regimes of traffic flow: free-flow and congestion. During the free-flow regime, vehicles travel at their desired

speed with no influence of the surrounding vehicles. During the congested regime, vehicles will not be able to drive at desired speed, influenced by the speeds of surrounding vehicles. Triangular FD is the simplest model consistent with empirical data and defined by three observable parameters: free-flow speed *u*, congested wave speed *w* and jam density κ.

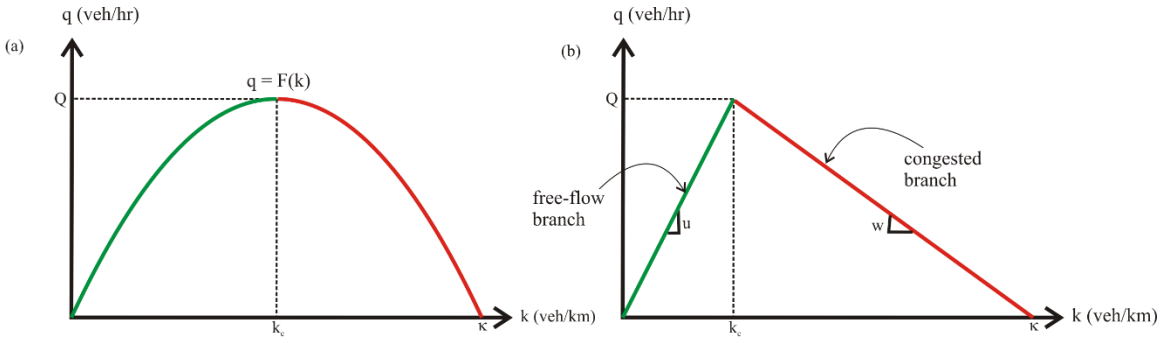

Figure 2: Fundamental Diagram a) Greenshields b) Triangular

The next section will briefly present the Newell-Daganzo merge model used in this study.

3.2. Merge Model

The Newell-Daganzo merge model (N-D model) [61] is one of the popular models empirically verified [69], and describes the merge process between two competing stream to provide the quantities of flows entering the merge. Let the two competing streams have capacities $\mu_1$and μ_2, demands λ_1 and λ_2, nominal flows $q_1$and q_2, and the merge ratio $p = q_2/q_1$. Let the merge capacity is μ_3 (see Figure 3a). The N-D model states that when both the competing streams are congested, the ratio of nominal flows, p, is a constant irrespective of their demands.

The N-D model represented in Figure 3b shows the flows on competing streams along the axes and the merge capacity, μ_3, indicated on both axes. The figure shows that

there are four sections where flows from competing streams are governed differently to determine the flows that can enter the merge. In section I, the combined flow from the streams is below the merge capacity. Therefore, merge flows are λ_1 and λ_2 and neither stream gets congested. In the rest of the three sections, the combined flow from the streams is greater than the merge capacity and hence there will be congestion on one or both streams. In section II, the demand on stream 1 (stream 2) is less (more) than its nominal flow. Therefore, the merge flows are λ_1 and $(\mu_3 - \lambda_1)$ and the stream 2 gets congested. In section III, the demand on stream 2 (stream 1) is less (more) than its nominal flow. Therefore, the merge flows are λ_2 and $(\mu_3 - \lambda_2)$ and the stream 1 gets congested. In section IV, the demand on both the streams is greater than their nominal flows. Therefore, the merge flows are given by q_1 and $q_2$and both streams get congested. Note than if the merge capacity falls below μ_3, due to downstream congestion, the merge capacity line is shifted to the left correspondingly.

One of the commonly used models for determining nominal flows is the *zipper* rule that gives equal priority to the shoulder lane and the ramp. Assuming *n* lanes on the freeway, the nominal flows for freeway and ramp are given by $\frac{n-0.5}{n}\mu_3$ and $\frac{0.5}{n}\mu_3$.

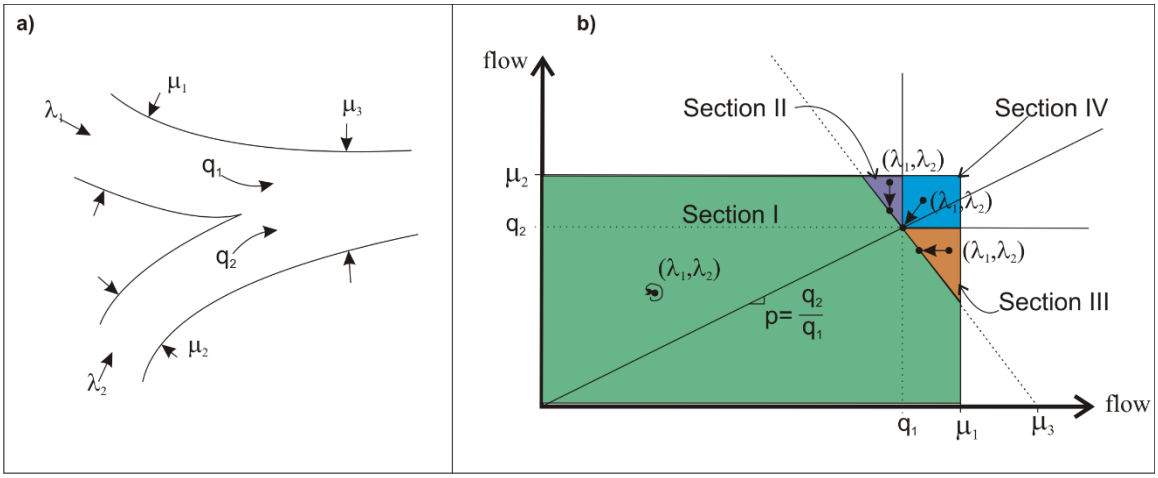

Figure 3: Newell-Daganzo Merge Model

The next section will describe the traffic flow models used in this study and their numerical solution methods.

3.3. Traffic Flow Models

3.2.1. First-Order Models

Traffic flow models can be divided as macroscopic, mesoscopic, and microscopic models depending on the level of details used to represent the traffic behavior. Macroscopic models treat traffic flow as continuum and use aggregate values for traffic variables to model traffic dynamics. Mesoscopic models increase on the level of detail to group vehicles with similar characteristics and use macroscopic traffic flow characteristics to model traffic behavior. Microscopic models provide the most detail by modeling lateral (car-following) and longitudinal (lane-changing) movement of individual vehicles.

The literature shows that macroscopic models in the state-space form are computationally efficient and most appropriate for solving traffic control problems [70], [71]. Macroscopic models are categorized into first-order, second-order, and higher-order

models depending on their complexity. The most popular first-order model is the Lighthill-Witham-Richards (LWR) model, also known as the KW model [72], [73]. The KW model is a scalar conservation law for the density, $k(t,x)$, of vehicles on a road at time t and location x, supplemented with a FD that relates the flow, $q(t,x)$, and density $k(t,x)$. For a road with net lateral inflow $g(k,t,x)$ from the entrances or exits the conservation law becomes:

$$\frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = g \qquad (5)$$

The equation *(5)* belongs to the family of first-order hyperbolic Partial Differential Equations (PDE). Generally, solutions to PDEs need initial and/or boundary conditions. Additionally, one needs to impose criteria for uniqueness and stability. In the case of KW model, unique and stable solution for a given initial or boundary value problem can be obtained by imposing the entropy condition [74]; i.e., a solution that maximizes the flow [75].

3.2.1.1. Godunov Scheme

Effective numerical solutions to the KW model started with the seminal work of Godunov [76]. In the area of transportation, this method is popularly known as Cell Transmission Model (CTM) [77], [78]. To illustrate the CTM steps for the KW model, assume a discrete grid with spatial and temporal discretization of Δx and Δt related by the *Courant–Friedrichs–Lewy* condition [79], which ensures that the analytical domain of dependence is contained within the numerical domain of dependence:

$$\frac{\Delta x}{\Delta t} \geq \tilde{C} \qquad (6)$$

where $\tilde{C} \triangleq \max(\frac{\partial F(k)}{\partial k})$ is the maximum characteristic speed of the FD. The consequence of this condition is that the wave stays within the cell during the computation time step.

Let the density in the cell *n* at time *i* be given by k_n^i and the *Cauchy* problem as:

$$k_x^0 = \begin{cases} k_u, & x = n \\ k_d, & x = n+1 \end{cases} \qquad (7)$$

where $k_u$and k_d are the densities in the two adjacent cells (upstream and downstream). (7) is called the *Riemann* problem in the classical PDE literature. Solutions to the *Riemann* problems are not trivial. But, for the case of KW model which has a concave FD, the solution becomes simple. Using discretization, the conservation equation *(5)* becomes:

$$k_n^{i+1} = k_n^i + \frac{\Delta t}{\Delta x}\left(q(k_{n-1}^i, k_n^i) - q(k_n^i, k_{n+1}^i)\right) \qquad (8)$$

Daganzo [77] showed that the flow *q* is given by the minimum of the demand and supply functions:

$$q(k_u, k_d) = \min(\lambda(k_u), \mu(k_d)) \qquad (9)$$

Note that both $\lambda(.)$ and $\mu(.)$ are continuous monotonic functions as shown in Figure 4 and *Q* is the capacity of the road. Then:

$$\lambda(k_u) = \min(k_u.u, Q) \qquad (10)$$

$$\mu(k_d) = \min(w.(\kappa - k_d), Q) \qquad (11)$$

One of the prominent features of the entropy solution is the propagation of deceleration waves (when the upstream density is smaller than the downstream density, $k_u < k_d$), given by the *Rankine-Hugoniot* condition (see Figure 4):

$$s = \frac{q(k_u) - q(k_d)}{k_u - k_d} \qquad (12)$$

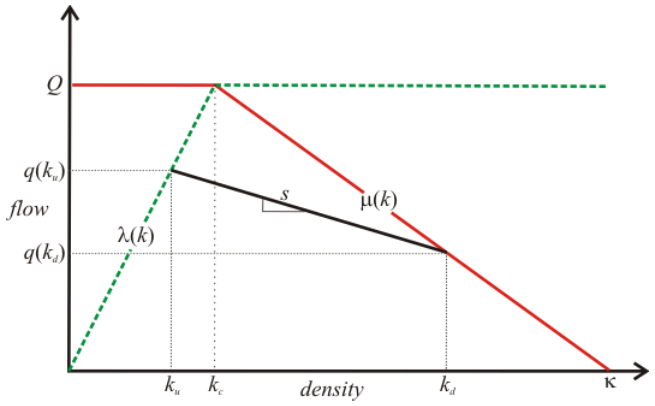

Figure 4: Demand and Supply functions

Empirical studies indicated that factors such as traffic conditions, weather, incidents, driving characteristics etc. affect the FD parameters [80]–[83]. Especially disruptive lane changes during the onset of congestion reduce capacity by as much as 10-15% [84]–[86]. Since first-order models do not model this phenomenon (called *capacity drop*), some researchers developed variations to the merge model [87], [88], FD [89]–[91] and demand functions [75], [92] to explicitly incorporate capacity drop into the numerical solutions (e.g. see Figure 5).

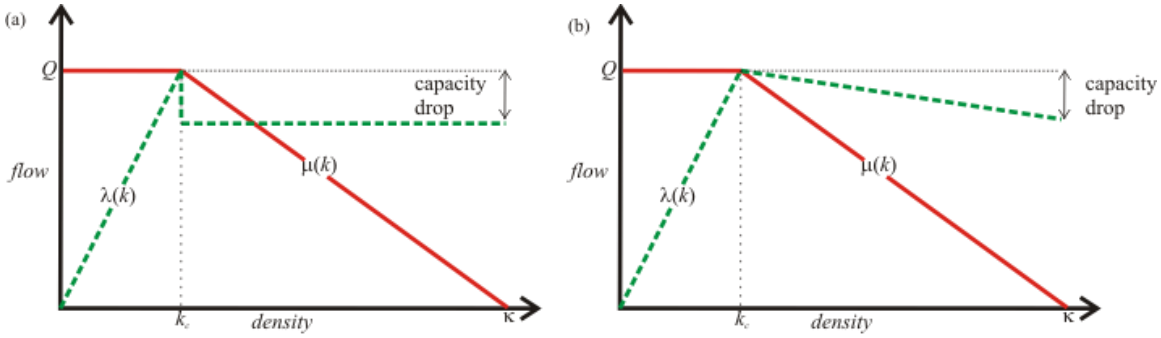

Figure 5: Sample demand functions a) Hadiuzzaman et.al. [92] b) Lebacque [75]

The demand function in Figure 5a indicates that the capacity drop occurs at critical density, k_c, which is a close representation of the reality. However, it makes the demand

function discontinuous at k_c:

$$\lambda(k) = \begin{cases} k.u, & k \leq k_c \\ Q(1 - c_d), & k > k_c \end{cases} \quad (13)$$

where c_d is the percentage capacity drop.

The demand function in Figure 5b is continuous and monotonously decreasing (capacity drop monotonously increasing) during the congestion regime:

$$\lambda(k) = \min(k.u, \frac{Q(Q - c_d Q - \kappa u + c_d.k.u)}{Q - \kappa u}) \quad (14)$$

Note that c_d=0 makes *(14)* equivalent to *(10)*.

This study uses *(14)* to explicitly incorporate capacity drop into the KW model.

3.2.1.2. Variational Theory

Unfortunately, Godunov scheme is known to introduce numerical errors for acceleration waves [93] since the characteristics do not always pass through the grid points.

Alternatively, Newell [94] proposed using Moskowitz functions to solve the KW model and conjectured that the numerical solution is the lower envelope of the curves translated from downstream and upstream. Daganzo [95], [96] proved the conjuncture using variational theory that enabled develop error-free solutions to the KW model.

Moskowitz functions, also popular known as *N-curves* or *cumulative count curves* are a powerful tool for analyzing the traffic dynamics on a system [97]. This function represents the evolution of cumulative number of vehicles that pass a reference point. Represented as $N(t, x)$, this function is tagged with a reference vehicle, monotonously increasing in time at any *x* and decreasing in space at any *t*.

When represented using *N-curves*, the conservation equation of the KW model becomes an identity. The FD reduces to the Hamilton-Jacobi equation:

$$\frac{\partial N}{\partial t} = F(-\frac{\partial N}{\partial x}) \qquad (15)$$

where $\frac{\partial N}{\partial t}$ and $-\frac{\partial N}{\partial x}$ represent the flow and density respectively.

For homogenous highways, $N(t,x)$ can be solved using Newell's minimum principle [94]. However, to solve inhomogeneous highways and moving bottlenecks, Daganzo [95], [98] used variational theory to show that:

$$N(P) = \inf_{B \in \beta_P} \{N(B) + \Delta_{BP}\} \qquad (16)$$

where *P* is the desired point in the time-space surface, *B* is any generic point on the *P*'s domain of influence, β_P, and $\Delta_{BP}$is the cost of the path from *B* to *P*, given by the maximum passing rate (see Figure 6). In essence, N value at P is the minimum of the N curves obtained from every point in its domain of dependence, which is exact during congestion.

The KW model is theoretically rigorous, numerically reliable, and computationally efficient. However, it does not model platoon diffusion, capacity drop, hysteresis, oscillations, and spontaneous onset of congestion etc. To overcome these limitation, researchers have developed second-order [99] and third-order models [100].

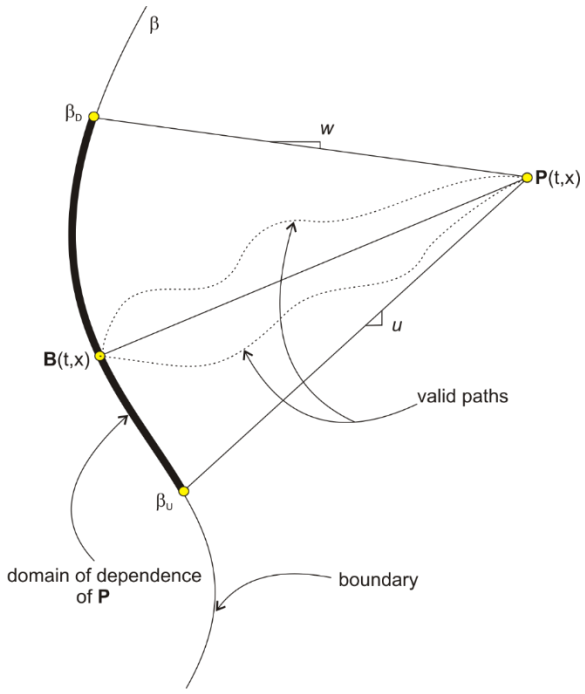

Figure 6: Visual representation of variational theory solution to the KW model

3.2.2. Second-Order Models

Second-order models supplement the first-order models with an extra equation for speed evolution and the third-order models supplement the second-order models with an extra equation for variance of velocity.

Payne [99] developed the first second-order model by supplementing the conservation equation *(5)* with the velocity (v) evolution equation:

$$\frac{\partial v}{\partial t} + v\,\frac{\partial v}{\partial x} + \frac{c_0^2}{k}\frac{\partial k}{\partial x} - \frac{V - v}{\tau} = 0 \qquad (17)$$

where $V(.)$ is the equilibrium velocity obtained from the FD, c_0 is the anticipation constant and τ is the relaxation constant. Equation *(17)* represents three mechanisms that influence the speed evolution: convection, anticipation and relaxation (given by the second, third and fourth terms of *(17)*). Convection represents the effects of acceleration

and deceleration on the average speeds of the segments. Anticipation takes into account driver's response to the density in the downstream segments. Finally, relaxation represents the drivers' acceleration and deceleration behavior.

Papageorgiou [52] developed a popular numerical methods for the second-order models :

$$k_n^{i+1} = k_n^i + \frac{\Delta t}{\Delta x}\left(q_{n-1}^i - q_n^i\right) \tag{18}$$

$$v_n^{i+1} = v_n^i + \frac{\Delta t}{\Delta x} v_n^i \left(v_{n-1}^i - v_n^i\right) + \frac{\Delta t}{\tau}\left(V\left(k_n^i\right) - v_n^i\right) - \frac{c_0^2 \Delta t}{\tau \Delta x} \frac{k_{n+1}^i - k_n^i}{k_n^i + \phi} = 0 \tag{19}$$

$$q_n^i = k_n^i \, v_n^i \tag{20}$$

where ϕ is the regularization parameter.

Note that the numerical solution, *(9)*, to the KW model is nonlinear in the flow calculation, but that for the second-order models it is linear *(19)* making them attractive for optimization problems. While system *(18)* - *(20)* model driver behavior more accurately, they have inherent weakness such as negative speeds, as identified in [101]. However, later researchers addressed the weaknesses and improved the second-order models making them relevant for modeling traffic [102], [103].

The next section will briefly present the optimal control theory methods for continuous time formulations.

3.4. Optimal Control Theory

Optimal control theory deals with methods to determine optimal control trajectory for a dynamic process, which are represented as ordinary differential equations (lumped systems) or partial differential equations (distributed parameter system).

The optimal control methods optimize a system for an objective function defined in terms of state and control variables and subject to constraints on the state and control variables. Optimal control theory methods for lumped systems is well developed [104]. Therefore, distributed parameter system models are typically reduced (using discretization) to use the lumped system optimal control methods. For continuous time systems calculus of variations methods are used and mathematical programming methods for discrete time systems as described below (interested readers are encouraged to read [104]):

Let the state of a system, control, and Lagrange multipliers are defined by a continuous vector functions $\boldsymbol{x}(t)$ (size: n x 1), $\boldsymbol{u}(t)$ (size: m x 1), and $\boldsymbol{p}(t)$ (size: n x 1), satisfies the state-space equation $\dot{\boldsymbol{x}}(t) = \boldsymbol{a}(x(t), u(t), t)$, initial and final times given by t_0 and t_f, and initial state be $\boldsymbol{x_0}$. The objective is to extremize:

$$J(\boldsymbol{u}) = h\big(\boldsymbol{x}(t_f), t_f\big) + \int_{t_0}^{t_f} g(\boldsymbol{x}(t), \boldsymbol{u}(t), t)dt \qquad (21)$$

where $h(.)$ and $g(.)$ are continuous functions.

As defined in the classical calculus of variations methods, the Hamiltonian is given by :

$$\boldsymbol{H}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{p}(t), t) \triangleq g(\boldsymbol{x}(t), \boldsymbol{u}(t), t) + \boldsymbol{p}^T(t).\boldsymbol{a}(x(t), u(t), t) \qquad (22)$$

Assuming the optimal state, control, and Lagrange multipliers are given by $x^*(t)$, $u^*(t)$, and $p^*(t)$, the necessary conditions for optimality are:

$$\dot{\boldsymbol{x}}^*(t) = \frac{\partial \boldsymbol{H}}{\partial \boldsymbol{p}}(\boldsymbol{x}^*(t), \boldsymbol{u}^*(t), \boldsymbol{p}^*(t), t\,) \qquad (23)$$

$$\dot{\boldsymbol{p}}^*(t) = -\frac{\partial \boldsymbol{H}}{\partial \boldsymbol{x}}(\boldsymbol{x}^*(t), \boldsymbol{u}^*(t), \boldsymbol{p}^*(t), t) \tag{24}$$

$$0 = \frac{\partial \boldsymbol{H}}{\partial \boldsymbol{u}}(\boldsymbol{x}^*(t), \boldsymbol{u}^*(t), \boldsymbol{p}^*(t), t) \tag{25}$$

$$\left[\frac{\partial h}{\partial \boldsymbol{x}}\left(\boldsymbol{x}^*(t_f), t_f\right) - \boldsymbol{p}^*(t_f)\right]^T . \delta \boldsymbol{x_f} + \left[\boldsymbol{H}\left(\boldsymbol{x}^*(t_f), \boldsymbol{u}^*(t_f), \boldsymbol{p}^*(t_f), t_f\right) + \frac{\partial h}{\partial t}\left(\boldsymbol{x}^*(t_f), t_f\right)\right] . \delta t_f = 0 \tag{26}$$

Note that *(23)* and *(24)* are a set of 2*n* first-order differential equations of the state and costate equations and *(25)* is a set of *m* algebraic relations which must be satisfied during $[t_0, t_f]$. Therefore, the recipe for the optimal control solution is:

1) Integrate the state and costate equations *(23)* and *(24)*
2) Evaluate the 2*n* constants of integration using $x^*(t_0) = x_0$ and eqn *(26)*, depending on how t_f and $x(t_f)$ are defined. Table 5-1 in [104] defines the solutions for different cases of boundary conditions.

In problems with linear state dynamics and quadratic performance it is possible to obtain the optimal control law by numerically integrating a matrix differential equations *(23)* - *(26)*. In general, however, this approach leads to a nonlinear two-point boundary-value problem that cannot be solved analytically.

To overcome this, the lumped systems are further reduced (time discretized) and solved using mathematical programming techniques such as steepest descent, variation of extremals, quasilinearization, gradient projection, etc.

This chapter presented a brief background on the merge model, traffic flow models and the optimal control methods, which are critical for this study. The next

chapter will present the optimality conditions for the case of unlimited ramp storage and continuous ramp formulation for the case of limited storage.

CHAPTER 4 ANALYTICAL FORMULATIONS

This chapter begins with a study of the optimality conditions for the case of unlimited storage on the on-ramps. We show that for this case, discrete ramp modeling using Moskowitz functions is sufficient to gain insights into the optimality conditions. However, explicit system modeling and optimization are necessary for the case of limited ramp storage. Towards this end, the second part of this chapter presents the continuum formulation of the optimal control problem to show how the nonlinearity in the FD renders it infeasible to solve analytically.

4.1. Discrete Ramps with Unlimited Queue

Consider a long freeway-ramps system with n on-ramps and m off-ramps as shown in Figure 7. Let the cumulative arrivals and departures for the ramps be $A_i(t)\ \forall\, i = 1, 2, \cdots\ n$ and $D_j(t)\ \forall\, j = 1, 2, \cdots\ m$ and for the freeway $A_f(t)$ and $(D_f(t))$.

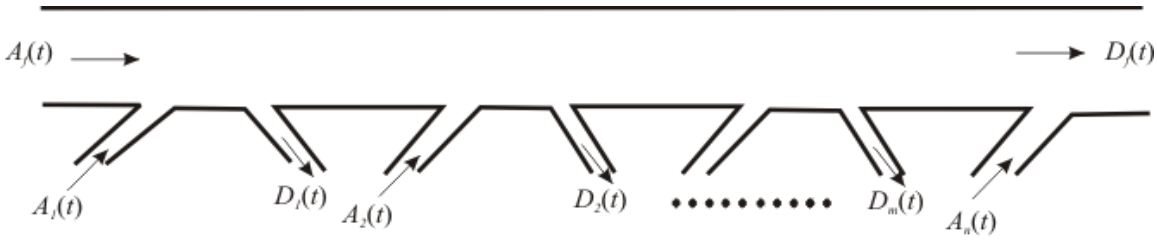

Figure 7: Corridor arrival and departure N-curves

The total delay for the vehicles during time *T* is given by (see Figure 8):

$$\int_0^T \big(A(s) - D(s)\big)\, ds \qquad (27)$$

where $A(s) = A_F(.) + \sum_{i=1}^{n} A_i(.)$ and $D(s) = D_F(.) + \sum_{i=1}^{m} D_i(.)$ are the aggregated

virtual arrivals and departures for the system. Minimizing *(27)* implies maximizing the second term because the first term of *(27)* is not controlled, i.e.:

$$max \int_0^T D(s)\, ds \triangleq max \left\{ \int_0^T D_F(.) + \int_0^T D_1(.) + \cdots\cdots \int_0^T D_m(.) \right\} \qquad (28)$$

Eqn. *(28)* is fundamental to deriving important insights for different scenarios as presented below. Alternatively, identical conclusions can be obtained when density and flows are used as state variables instead of $N(t)$, as shown in Appendix A.

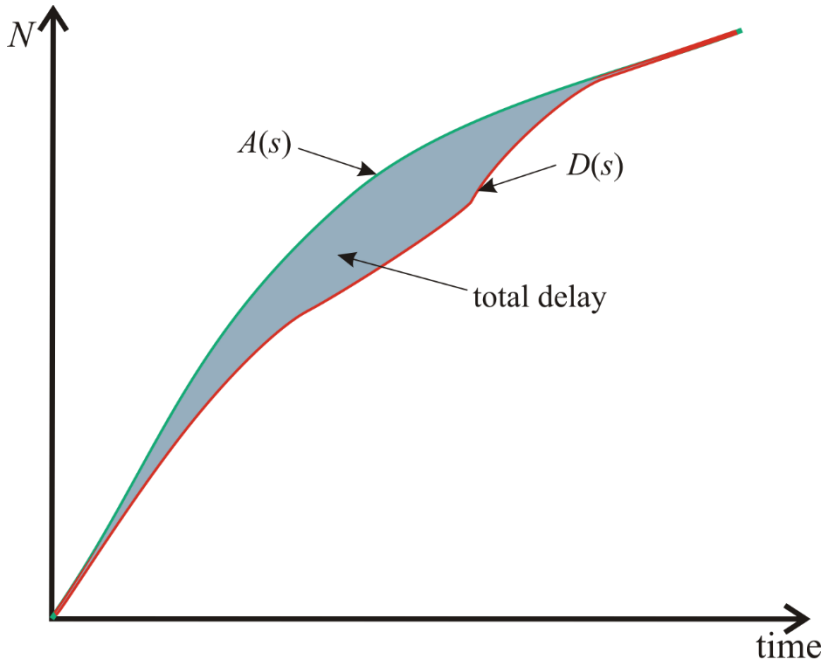

Figure 8: Arrival and Departure Curves

Consider the case of a corridor whose exit flows are defined as a proportion of freeway flow. Also, assume that there is unlimited storage on each on-ramp.

Insight 1: *System-wide optimal metering can be achieved with isolated metering*.

Since the on-ramps have unlimited storage, there is no lower bound on the allowable metering rate. Therefore, *each* off-ramp outflow can be maximized with an associated upstream ramp meter that maximizes freeway flow. Therefore, *(28)* becomes:

$$max \int_0^T D_F(.) + max \int_0^T D_1(.) + max \int_0^T D_2(.) + \cdots max \int_0^T D_m(.) \qquad (29)$$

Thus, for the case of unlimited storage, the system optimal control problem can be broken down into an isolated control problem. Therefore, the optimality conditions for the case of unlimited storage can be defined as:

Only at the onset of capacity conditions should the on-ramps that experience congestion be metered at $\beta_j . Q$ to maintain freeway at critical density (β_j the exit proportion at the upstream off-ramp and Q is the freeway capacity).

Justification: The concave nature of the FD implies that $k_c = \underset{k}{argmax}(F(k))$. Therefore, the freeway operated at critical density k_c (see Figure 2) is the global optimality. Therefore, based on *insight 1*, each on-ramp maintaining critical density k_c on the freeway is the system optimal control. Controlling the on-ramps before the freeway reaches capacity will delay the onset of capacity conditions and hence increase total delay. Therefore, ramp metering should be initiated only at the onset of capacity conditions. Also, since the system-optimal control problem is equivalent to local optimal control problem, only the ramps that potentially get congested need to be controlled. The ramps that do not get congested will incur delay with any control. Once the freeway is at critical density, the vehicle conservation between off-ramp flow and on-ramp flow indicates that the metering rates should be equivalent to the exit flow i.e. $\beta_j Q$.

Note that the above optimality conditions are consistent with the popular heuristic notion that the ramp meters should aim to maintain critical density on the freeways. Also, the demand capacity method *(1)* provides a metering rate of $\beta_j.Q$ at the onset of capacity conditions. These results indicate that the heuristic theories proposed based on experience are truly optima under the assumptions of unlimited on-ramp storage and exogenous outflow proportions.

Insight 2: *For the many-to-one problem when capacity drop is ignored, on-ramp control is not needed.*

For this case, *(29)* reduces to:

$$max\left\{\int_0^T D_F(.)\right\} \qquad (30)$$

If there is enough demand to create freeway congestion and the destination always operates at capacity, total delay is invariant to the control strategy used at the on-ramps because bottleneck flow cannot be increased with metering.

Now we go back to *(28)* and try to analyze for the case of limited storage. One of the consequences of limited on-ramp storage is the potential for queue spillback, which is not acceptable for most operators. Therefore, there will be lower bounds on the metering rate and hence it is not possible to maintain critical density on the freeway. Then, equation *(28)* has infinitely many solutions and finding the optimal solution requires optimization methods.

Conjecture: *For the case of limited storage, minimizing total delay is achieved by maximizing freeway flow near ramps with the highest exit proportions $\beta_j s$.*

Using exit proportions, $\beta_j(t)\ \forall\ j = 1, 2, \cdots\ m$, *(28)* becomes:

$$max\left\{\int_0^T D_F(.) + \int_0^T \beta_1(.).q_1(.) + \int_0^T \beta_2(.).q_2(.) + \cdots \int_0^T \beta_m(.).q_m(.)\right\} \qquad (31)$$

where, $q_j(t)\ \forall\ j = 1, 2, \cdots\ m$ are the freeways flows at each ramp. While the β_js are independent across ramps, the q_js are correlated. The assumption that the system evolves slowly (proposed by Daganzo [105], which means that the ratio of freeway flow to the ramp flows is high and changes in freeway flows does not abruptly change freeway dynamics) mean that each term of *(31)* can be assumed independent. Thus, maxima of *(31)* may be achieved by metering in such a way that the freeway flow is maximized at the exits with high β_js. Incidentally, identical results are obtained by Daganzo [105] using variational methods for the case of maximizing flows on congested cities.

This section presented the optimality conditions for the case of unlimited storage using Moskowitz functions. However, as mentioned above, for the case of limited storage, explicit system modeling and optimization will be necessary. Towards this end, the next section will describe the continuum ramp formulation under queue constraint.

4.2. Continuum Ramp Formulation under Queue Constraint

This study models the ramps and freeway using the KW model, with the control governing the number of vehicles that transfer between ramps and freeway. The control appears in the source term. Therefore, the optimal control problem of the freeway-ramps system is a coupled PDE system with distributed control as shown below:

Consider a long homogenous freeway corridor and on-ramps of length L and l respectively. If the number of entrances and exits are large, the inflow and outflow can be represented continuous in space and time. Let the density be $k(t, x, y)$, where x and y represent the distance axes along freeway and on-ramps. Assuming both freeway and

ramps are governed by the same FD, and the exit flows are a proportion, $\beta(t,x)$, of the freeway flow, the freeway-ramp system with the control trajectory, $u(t,x)$, becomes:

$$\frac{\partial k(t,x,l)}{\partial t} + \dot{F}(k)\frac{\partial k(t,x,l)}{\partial x} = \min\left(u(t,x), \varphi^{+}\left(k(t,x,l)\right)\right) - \beta(t,x).F(k(t,x,l)) \quad (32)$$

$$\frac{\partial k(t,x,y)}{\partial t} + \dot{F}(k)\frac{\partial k(t,x,y)}{\partial y} = 0 \quad (33)$$

subject to the boundary conditions:

$$\begin{gathered} k(t,x,0) = \alpha(t,x), \\ k(t,x,l) = min(u(t,x), \varphi^{+}(k(t,x,l)))), \\ k(t,0,l) = 0, \\ k(t,L,l) = 0 \end{gathered} \quad (34)$$

where $\varphi^{+}$is the flow from the N-D merge model [61].

This study uses total time spent (TTS) as the objective function, which is given by:

$$min \int_0^X \int_0^Y \int_0^T k(t,x,y)\, dt\, dy\, dx \quad (35)$$

Note that minimizing TTS is the same as minimizing total delay because the TTS includes both free-flow travel time and delay, and the former is invariant to the control.

Systems *(32)-(33)* and *(35)* belong to the class of distributed parameter systems that are difficult to solve analytically and generally addressed using model reduction by spatial discretization to make it a lumped system. To see this, let Δx and Δy be discrete cell sizes, *i* and *j* be the discretization indices along freeway and ramps, and assume forward differences to represent spatial derivative. Then *(32)-(33)* and *(35)* becomes:

$$\dot{k_\iota}(t,l) = -\dot{F}(k)\left(\frac{k_{i+1}(t,l) - k_i(t,l)}{\Delta x}\right) + \tag{36}$$

$$min(u_i(t), \varphi^+(k_i(t,l), k_{il}(t))) - \beta_i(t).F(k_i)$$

$$\dot{k_{\iota J}}(t) = -\dot{F}(k)\left(\frac{k_{ij+1}(t) - k_{ij}(t)}{\Delta y}\right) \tag{37}$$

$$min \int_0^T (\sum_i (k_i + \sum_j k_{ij}))\, dt \tag{38}$$

Comparing the state-space representation *(36)-(38)* with that in 3.4,

$$\dot{x}(t) = [\dot{k_\iota}(t,l), \dot{k_{\iota J}}(t)]^T \tag{39}$$

$$h(.) = 0 \tag{40}$$

$$g(.) = \sum_i (k_i + \sum_j k_{ij}) \tag{41}$$

$$a(.) = [-\dot{F}(k)\left(\frac{k_{i+1}(t,l) - k_i(t,l)}{\Delta x}\right)$$

$$+ \min(u_i(t), \varphi^+(k_i(t,l), k_{il}(t))) - \beta_i(t).F(k_i), \tag{42}$$

$$-\dot{F}(k)\left(\frac{k_{ij+1}(t) - k_{ij}(t)}{\Delta y}\right)]^T$$

Notice that $a(.)$ is nonlinear due to the nonlinear FD (piecewise linear) that has state dependent characteristic speeds, *min* function, and nonlinear merge model that gives unique value of the flow depending on which of the 4 section, the flows are located. Thus, the state-space representation *(36)-(37)* is uniquely nonlinear and does not fit generic formulations that can be solved analytically using necessary conditions presented in 3.4. Therefore, this model is further reduced with time discretization as shown in the next chapter.

CHAPTER 5 DISCRETE TIME-SPACE FORMULATIONS

Since the continuous time formulation in the previous chapter cannot be solved analytically, the first section of this chapter presents a discrete-time formulation of the freeway ramp system presented in (36)-*(38)*. The second section presents an error-free solution to the KW model with source term that enhance the quality of the numerical solution. Section 3 presents variations of the discrete-time formulations of section 1 to derive the LP formulation used in the rest of the study.

Let the time-space surface in section 4.2 be discretized into *n* cells along the freeway and *m* cells along the ramps, k and q (or f) be the density in veh/km and flow in veh/hr, β_i be the exit proportion in 1/km, and α_i be the entry demand in veh/hr-km, as shown in Figure 9.

For $i = 1,2, \dots n$, $j = 1,2, \dots m$, and $t = 1,2, \dots p$, the conservation equation for the freeway *(36)* in the time-discrete form is a modification of *(8*) due to the source term:

$$k_i^{t+1} = k_i^t + \frac{\Delta t}{\Delta x}\left((1 - \beta_i^t).f_i^t - f_{i+1}^t + q_i^t\right) \qquad (43)$$

The conservation equation for the ramps *(37)* is the same as that given in *(8*):

$$k_{ij}^{t+1} = k_{ij}^t + \frac{\Delta t}{\Delta x}\left(q_{ij+1}^t - q_{ij}^t\right) \qquad (44)$$

The flows, f^t and q^t for *(43)* and *(44)* are calculated using the Godunov scheme as follows:

$$f_i^{t+1} = min(\mu(k_{i+1}^t), \lambda(k_i^t)) - q_{i+1,1}^t \qquad (45)$$

$$q_{ij}^{t+1} = min(\mu(k_{ij-1}^t), \lambda(k_{ij}^t)) \qquad (46)$$

Note that *(45)* is an important modification to the classical solution *(9)* of the *Riemann* problem to account for the source term. This modification ensures that the flows

are in equilibrium because the total flow entering the cell is strictly based on the density in the previous time step. This ensures that the flows fall on the FD boundary.

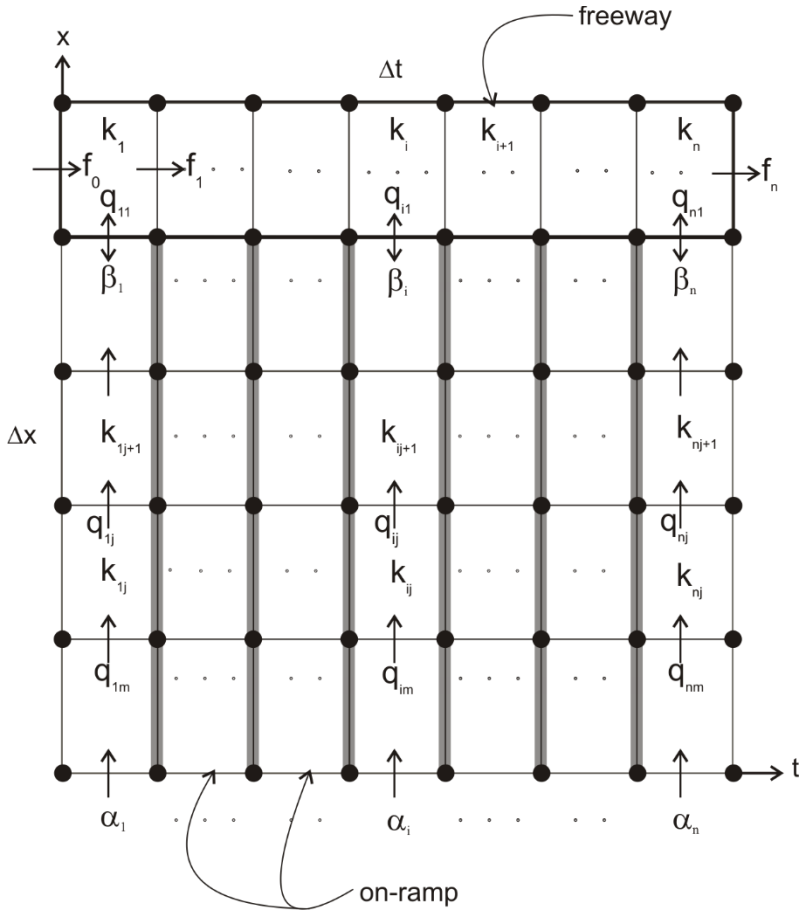

Figure 9: Time-space grid for discrete formulation

Assuming the ramp meters are located at the downstream boundary of the ramps, with a metering rate r_i^t, the merge flows are given by:

$$q_{i1}^{t+1} = min(r_i^t, \lambda(k_{i1}^t).min\left(1, \frac{\mu(k_{i+1}^t)}{\lambda(k_i^t)}\right)) \qquad (47)$$

where the second term represents the merge model in [106] which is proved to be equivalent to the N-D merge model [61] in the limit. Note that the merge model in *(47)* is

highly nonlinear with the ratio of two *min* functions, which we show later how it renders the optimal control problem difficult to solve.

Note that q_{im+1}^t is α_i^t, and f_0, α_i, and β_i, are known a priori. The queue constraint is posed as:

$$k_{im}^t \leq \kappa - \frac{\alpha_i^t}{w} \tag{48}$$

Note that equation *(48)* ensures that the supply capacity into the upstream most cell is at least the ramp flow.

Finally, the objective function *(35)* is a simple transformation:

$$min \sum_{t=1}^{p} \sum_{j=1}^{m} \sum_{i=1}^{n} (k_i^t + k_{ij}^t) \tag{49}$$

Several extensions to the nonlinear system *(43)-(49)* will be presented in section 5.3 to make the problem solvable in reasonable time. Before we do that, the next section presents an error-free solution to the KW model with a source term since the classical Godunov scheme induces errors during congestion.

5.1. Exact Godunov Scheme

The errors in the classical Godunov scheme arose because of the presence of multiple characteristic speeds that do not always pass through the grid points. To overcome this, we exploit the Daganzo's variational formulation [95], [96], [98] to develop an Exact Godunov (EG) scheme. The idea is to use the N-value from the location where the congestion characteristic passes through the grid point (exact value) to determine the N value at the desired grid point.

Let the time-space surface is discretized in Δt time units and Δx space units, a triangular FD with $\frac{u}{w} = p$, an integer, q_i be the grid flow, f_i be the source term, and N_i be the *N*-value at grid point *i*, as shown in Figure 10.

Daganzo's variational theory [98] shows that for the special case of a homogenous highway governed by a triangular FD, the cost function is linear and hence, all the valid paths between any two points have the same cost. This conclusion implies that the *N*-values are conserved along all paths between two points. Exploiting this, the conservation of vehicles along path 2 and path 1 from Figure 10 is:

$$N_D + k_D.\Delta x \ + \Delta t \ . \sum_{i=0}^{p} q_i + \Delta t \ . \sum_{i=0}^{p-1} f_i = N_D + \kappa.\Delta x \qquad (50)$$

Solving *(50)* for q_p,

$$q_p = \frac{\Delta x}{\Delta t}(\kappa - k_D) - \sum_{i=0}^{p-1}(q_i + f_i) \qquad (51)$$

Note that during free-flow, q_p calculated from N_D, which is exact, will be higher than that calculated from N_U. However, during congestion, the opposite is true and the exact value from $N_D$results in an error-free solution.

Incorporating *(51)* into classical Godunov scheme equation *(9)*, the EG scheme is represented as:

$$q^t(.) = \min(\lambda(k_u^t), \frac{\Delta x}{\Delta t}\left(\kappa - k_d^{t-p}\right) - \sum_{i=1}^{p-1}(q^{t-i} + f^{t-i})) \qquad (52)$$

where k_d^{t-p}represents the density of the downstream cell *d*, *p* time steps earlier.

Note that the EG scheme requires memory as one needs to track the grid flow and net inflow for the last *p* time steps and the downstream cell density *p* time steps earlier.

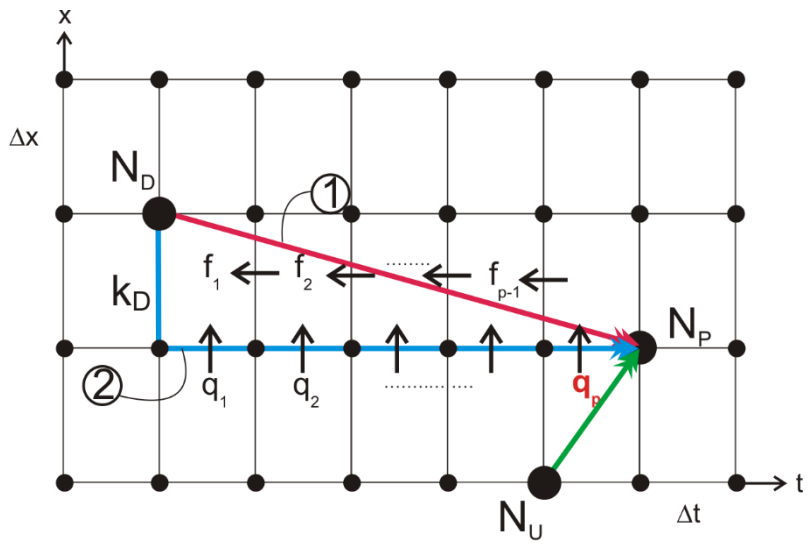

Figure 10: Flow calculation using exact Godunov method

Figure 11 presents the time-space density contours which show that unlike the classical Godunov scheme, the EG scheme maintains a distinct shock propagation.

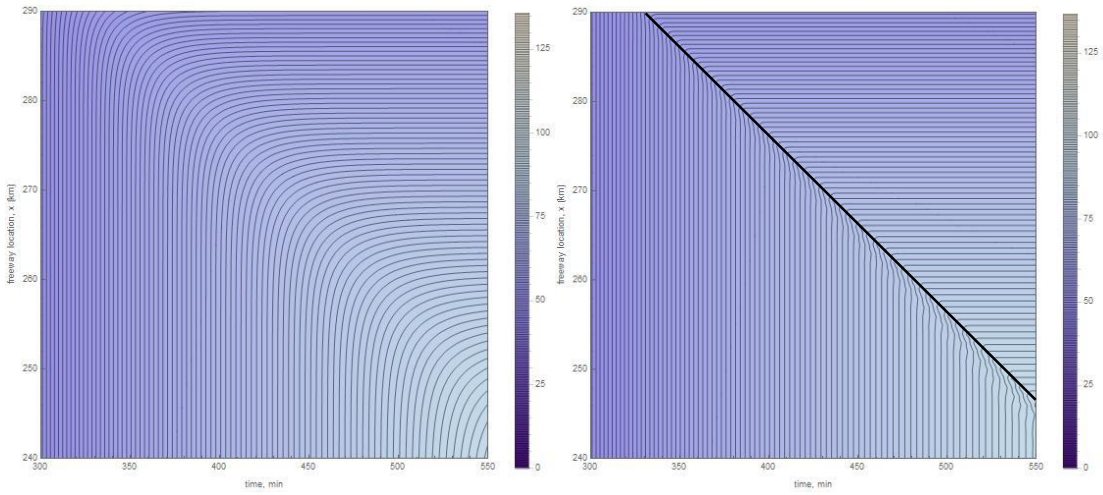

Figure 11: Shock propagation with Classical Godunov and Exact Godunov methods

The nature of nonlinearity in equations (45)-(47), with *min* function and merge model, makes it less suitable to solve efficiently with gradient-based methods that use

derivatives (gradients and Hessians). Such problems can be solved using direct search methods or metaheuristics as described in the next section.

5.2. Simulation-Based Optimization

Simulation-based optimization (SBO) is widely used for traffic problems [107]–[111]. These methods are popularly used for parameter optimization problems with discrete search space since they are known to get out of local optima and converge to global optima with appropriate algorithmic parameter values [112]–[114].

This study employs a SOB framework (shown in Figure 12) in Mathematica© 9.0.1 [115] and uses the inbuilt Differential Evolution (DE) optimization method, which is found to converge faster than other methods of its class; Nelder-Mead, Simulated Annealing, and Random Search methods. The framework contains two modules; the macrosimulation-based on the discrete formulation *(43)-(49)* and the DE optimizer. The process is iterative with the optimizer sending a combination of control values and the microsimulation returning the objective function value until the optimal solution is found.

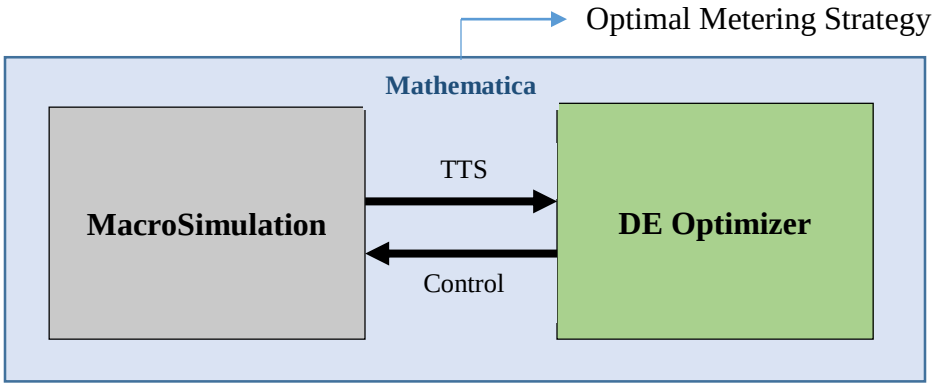

Figure 12: Simulation-based optimization framework

Since metaheuristic methods are known to be computationally expensive and converge slowly [109], [116], the simulation execution speed becomes critical. To gain efficiency in simulation speed, this study uses functional programming style [117], as against the intuitive imperative programing style.

Imperative programming style uses statements to explicitly indicate *how* to do something; e.g. array manipulation using this style involves iterating through each element of the array and performing operations on the element. Functional programing style uses expressions to tell *what* you would like to happen, and let the compiler figure out *how* to do it; e.g. for array manipulation, the compiler evaluates every element of the array simultaneously and avoids searching by indices Though functional programming style is not intuitive and often confusing, it requires orders of magnitude less code and more speed than imperative programming, which results in significant benefits as the *system* gets bigger.

In addition to the style, other optional choices available in Mathematica (*inline*, *compile*, *multi-threading*, and *parallel processing*, etc.) are exploited. The *inline* option enables defining efficient custom functions to be called from other efficient functions. The *compile* option generates a sequence of simple instructions for quick evaluation by a virtual machine by storing the intermediate results in the temporary memory.

Using these methods, the speed of the simulation is increased significantly (by over 215 times compared to imperative programming style). For example, a 3600 cell grid with 750 time steps is simulated in less than 0.8 seconds including the initiation, execution, communication, and termination time.

DE is a popular stochastic, robust, and population-based evolutionary algorithm developed by Storn and Price [118]. This algorithm requires values for three parameters: cross probability, population size, and scaling factor. Cross probability is used to generate new individuals from their parents, population size determines the number of individuals in each generation, and scaling factor is used to create a hybrid individual from existing individuals.

Researchers have noted that optimal solution is not always guaranteed with DE, but a careful selection of DE parameter values will increase the chances of success. Therefore, parameters values are determined using *Meta-Optimization* [109]. This method involves using hierarchical optimization on a smaller solution space with a known global optima (obtained via exhaustive search) to determine optimal DE parameter values based on their rate of convergence to the known global optima. Based on *Meta-Optimization,* the optimal parameter values for cross probability, population size, and scaling factor were found to be 0.8, 100,000, and 0.4. Additionally, no upper bound was specified on the maximum number of iterations or generations.

The SBO framework shown in Figure 12 is executed on the Georgia Institute of Technology's PACE high performance computing clusters [119].

Extensive evaluations revealed that SOB framework failed to converge in a reasonable amount of time when the number of decision variables and the corresponding search space was large. For a sample continuum ramp case of 90 freeway cells simulated for 750 time steps, assuming the metering rate changes every time step, there are 67,500 decision variables. Assuming metering rate $r \in [0, 2500]$, the size of the search space

becomes 168.75 million points (exhaustive search would need over 4 years to converge). Even with optimal DE parameters, the framework did not converge in 3 weeks.

Based on the insights gained from 0, the metering is initiated only when freeway reached critical density, which reduced the number of decision variables to be optimized. Then, granularity for control values is increased; metering rate increases or decreases in the order of 100 veh/hr-km. Next, the granularity of the "control region" is increased, where ramp groups operated at the same rate at any given time. These strategies significantly reduced the number of decision variables and the search space.

However, the results indicated a large number of solutions with the same objective function value, which is an artifact of the granularity (there were over 40,000 control strategies within 1% of the "optimal" TTS value). Moreover, there was no distinct pattern for the solutions indicating potential convergence to local optima. Most importantly, the optimization results for the case of unlimited storage did not converge to the optimal solutions obtained in section 4.1. Therefore it was determined that the simulation-based framework may not be suitable for optimal control problems.

The next section will present variations of the nonlinear discrete time formulation in *(43)-(49)*. We show that the nonlinear programming formulation and mixed integer programing formulation did not yield results in reasonable amount of time. To overcome this, we relax the nonlinear problem to pose it as a linear programing problem.

5.3. LP Formulation

We fall back on the discrete formulation *(43)-(49)* and explore ways to use mathematical programming methods. The current formulation is cumbersome for the following reasons:

a) Since the solution to the KW models is nonlinear, it makes every flow variable nonlinear, resulting in a large pool of nonlinear variables.
b) Explicit ramp modeling under the continuum approximation framework has a large number of cells, which significantly increases the number of variables.

Several attempts to solve the NLP problem did not produce any feasible solution, even with a quadratic objective function. The problem is then reposed as a Mixed Integer Linear Programming problem by replacing the "*min*" function using standard techniques [120], [121] and solved using the commercial solver LINDO© [122]. Several combinations of options such as quadratic objective function, defined/undefined bounds, were tested which proved to be in vain.

To circumvent this, we use the techniques from Gomes and Horowitz [42] to develop LP formulation for this study. Gomes and Horowitz replace the min function in the Godunov scheme with inequalities based on insights from Papageorgiou [55] and Ziliaskopoulos [123]. However, the objective to minimize delay ensures that the flows take the upper bound.

The formulation used in this study differs from Gomes and Horowitz 's formulation in the following ways:

Explicit Modeling of On-ramps

The studies in the literature model on-ramps using point queue methods [1], [9], [25], [42], [124], which does not account for ramp dynamics. To overcome this, freeway equations *(45)* are supplemented with *(46)* to explicitly model traffic dynamics on the ramps. The results indicate that this formulation enables unveiling important relationships between control trajectory and ramp dynamics as shown in 6.1.

Incorporate Capacity Drop:

This study also extends the Gomes and Horowitz's formulation by using the Lebacque's demand function [75] to incorporate the capacity drop phenomenon, which is critical for ramp metering problems:

$$\lambda(k_u) = min(k_u.u, Q, \left(1 - \frac{w_d}{w}\right) Q + w_d.(\kappa - k_d)) \qquad (53)$$

where w_d is the slope of the demand curve in the congested regime; see Figure 5b.

Ramp Priority Assumption

For the purpose of this study, which aims to understand the optimality conditions, we relax the N-D merge model. We assume full priority to the ramp flows to ensure metered flows enter the freeway uninterrupted. Based on the optimality conditions, we make qualitative arguments for the influence of the merge model. We also derive bounds of the dimensionless variables, for which the full priority assumption is valid. Thus, *(47)* becomes:

$$q_{ij}^{t+1} = u_i^t \qquad (54)$$

Objective Function

Studies have optimized ramp meters for single ([1], [36], [54], [125]) or multiple ([42], [124]) performance measures. Unlike Gomes and Horowitz, we use TTS as the single performance measure. This is because maximizing flow is invariant to the control strategy when the system is emptied.

With the modifications and assumptions made above, the objective function and the constraints used in the rest of this study are:

Objective $$min \sum_{t=1}^{p}\sum_{j=1}^{m}\sum_{i=1}^{n}(k_i^t + k_{ij}^t)\,\Delta x\,\Delta y\,\Delta t$$

Subject to:

Freeway Conservation $$k_i^{t+1} = k_i^t + \frac{\Delta t}{\Delta x}\left((1-\beta_i^t).f_i^t - f_{i+1}^t + q_i^t\right)$$

Ramp Conservation $$k_{ij}^{t+1} = k_{ij}^t + \frac{\Delta t}{\Delta x}\left(q_{ij+1}^t - q_{ij}^t\right)$$

Freeway Flow $$f_i^{t+1} \le (\kappa - k_{i+1}^t)w - q_{i+1,1}^t$$

Freeway Flow $$f_i^{t+1} \le (k_i^t.u) - q_{i+1,1}^t$$

Freeway Flow $$f_i^{t+1} \le \frac{Q(Q - c_d Q - \kappa u + c_d.k_i^t.u)}{Q - \kappa u} - q_{i+1,1}^t$$

Freeway Flow $$f_i^{t+1} \le Q - q_{i+1,1}^t$$

Ramp Flow $$q_{ij}^{t+1} \le \left(\kappa - k_{ij-1}^t\right)w$$

Ramp Flow $$q_{ij}^{t+1} \le \left(k_{ij}^t.u\right)$$

Ramp Flow $$q_{ij}^{t+1} \le Q$$

Control $$q_{i1}^{t+1} = r_i^t$$

Queue constraint $$k_{im}^t \le \kappa - \frac{\alpha_i^t}{w}$$

This chapter presented a LP formulation used in the rest of the study. The next chapter presents the results obtained with solving the LP problem.

CHAPTER 6 ANALYSIS WITH LP FORMULATION

In Chapter 4, we use Moskowitz functions to derive optimality conditions for the case of unlimited storage. But such results are not possible to achieve for the case of limited storage. Therefore, we use the LP formulation developed in chapter 5 to obtain optimal control solutions using optimization. Towards this end, this chapter presents important insights into the qualitative aspects of the traffic dynamics during the optimal control, the optimality conditions, and the nature of optimal control.

We apply the classical Godunov scheme to solve the Riemann problem for the numerical example using the parameters shown in Table 1. The resulting LP formulation is solved using LINDO© 15.0.34 [122]. The main insight gained from the time-space optimal metering surface is that there exists 4 zones with distinct metering pattern. These zones (labeled A, B, C, D) are logically related to the corresponding state trajectory on the freeways and ramps. One zone (A) is always in free-flow, while rest have both free-flow and congested conditions during the peak period. Also, zones C and D exhibit the same metering pattern during peak period, but differ when the demand drops.

Table 1: Parameters for Numerical Experiment

Parameter	Δx (km)	Δt (hr)	κ (veh/km)	w (km/hr)	u (km/hr)	c_d	n	p	m
Value	0.1	0.001	150	20	100	10%	30	90	10

6.1. Monotonic demand

Two parameters that we study in this analysis are the ramp demand and exit proportion, whose monotonic demand patterns are studied. We study the following space varying patterns for α (x) and β(x):

- α (x) = β(x) = a - (constant)
- α (x) = β(x) = ax - (increasing)
- α (x) = β(x) = a(1-(x/L)) - (decreasing)

The above pattern indicates nine possible combinations for α (x), β(x). We begin the analysis with the case of spatially constant on-ramp demand $\alpha\ (x) = a$ veh/hr-km and exit proportion $\beta(x) = b$/km and show the resulting spatiotemporal optimal control trajectory in Figure 13.

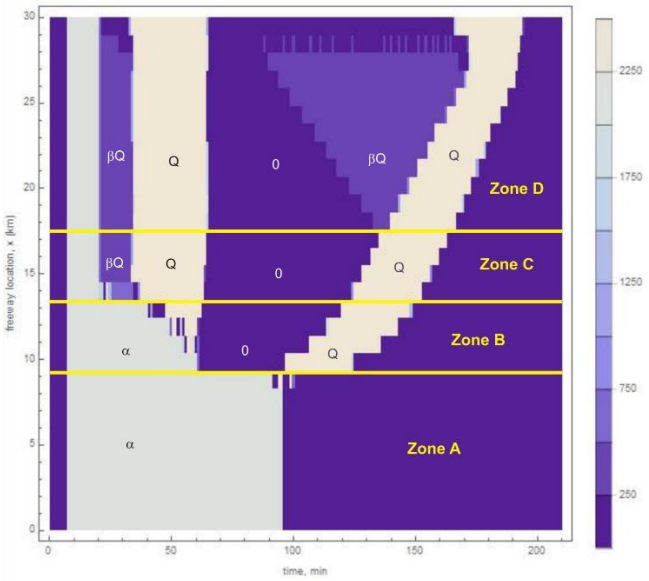

Figure 13: Time-space metering rate for a =2000 veh/hr-km and b =0.1/km

Figure 13 shows that there are four zones with distinct temporal metering patterns with the following metering rates:

- Zone A - α
- Zone B - α, 0, Q
- Zone C - α, βQ, Q, 0, Q
- Zone D - α, βQ, Q, 0, βQ, Q

These unique patterns are correlated with the freeway traffic dynamics as in Figure 14.

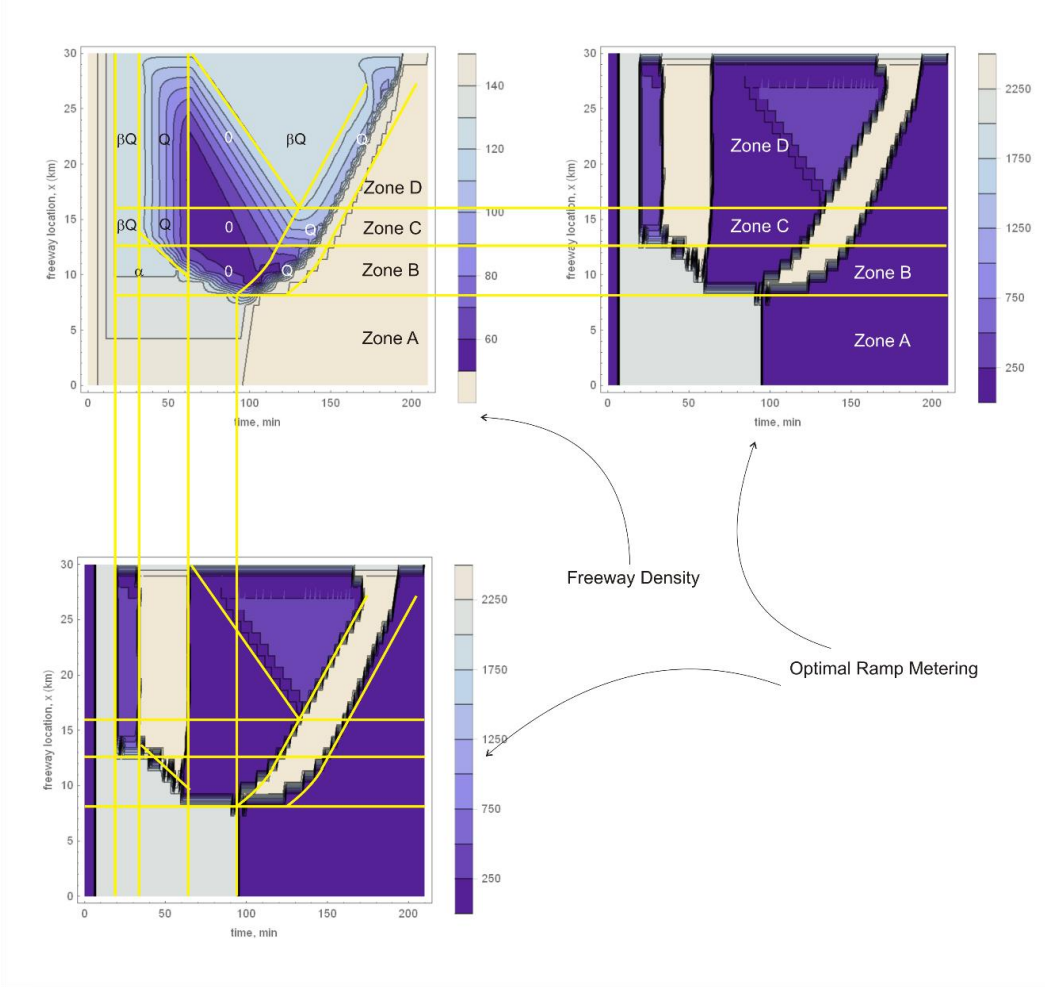

Figure 14: Relationship between optimal control strategy and state for a =2000 veh/hr-km and b =0.1/km

Figure 14 indicates that the metering patterns are correlated with the shock evolution and density evolution on the freeway. However, the queue propagation on the ramps is critical

to understand the correlations. The following subsections describe the unique characteristics of optimal metering for each zone with respect to the traffic dynamics on the ramps.

Zone A

Figure 14 shows that zone A corresponds to the transient "loading" of the freeway and the steady-state free-flow equilibrium. All its locations have the same metering pattern: rate equivalent to the demand until loading and drops to 0 (it should be interpreted as meter being turned off) when the demand is turned off; see Figure 15.

Insight: *Zone A should not be metered*. This validates the argument that ramps that are not congested should not be metered from is validated.

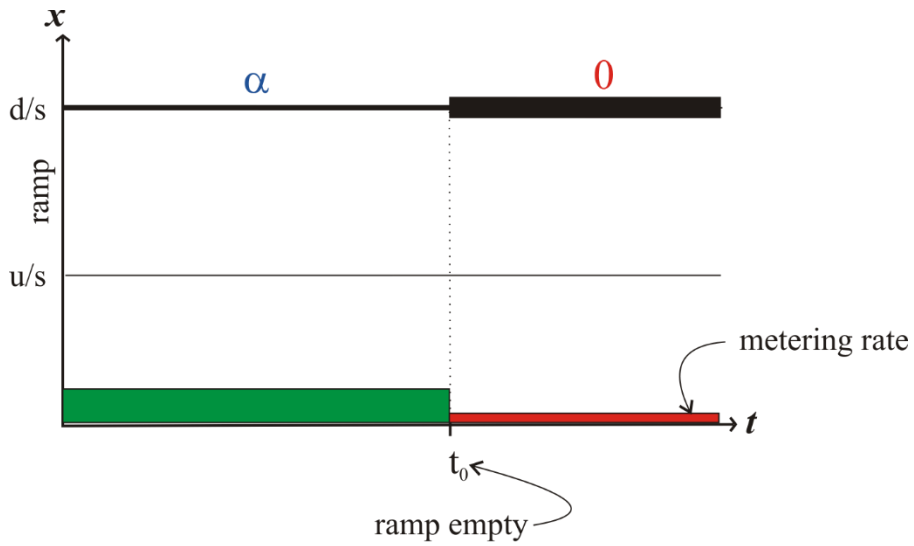

Figure 15: Optimal metering and ramp queues in Zone A

Zone B

Since it is clear that metering should be initiated only after congestion begins, we eliminate the "loading" period from the discussion. Figure 14 shows that Zone B corresponds to the steady-state free-flow equilibrium and then congestion. The results in

Figure 16 indicate that the metering is not activated until time t_1, time when the ramp is shut down such that the shock s_2 reaches the upstream boundary just when the ramp empties (see Figure 17). This ensures that the queue remains at the upstream boundary, even with ramp shutdown. Using the slope s_2 in Figure 17, the analytical solution for t_1 becomes:

$$t_1 = t_0 - \frac{l_r(n\kappa u - \alpha)}{\alpha u} \qquad (55)$$

where l_r is the length of the ramp.

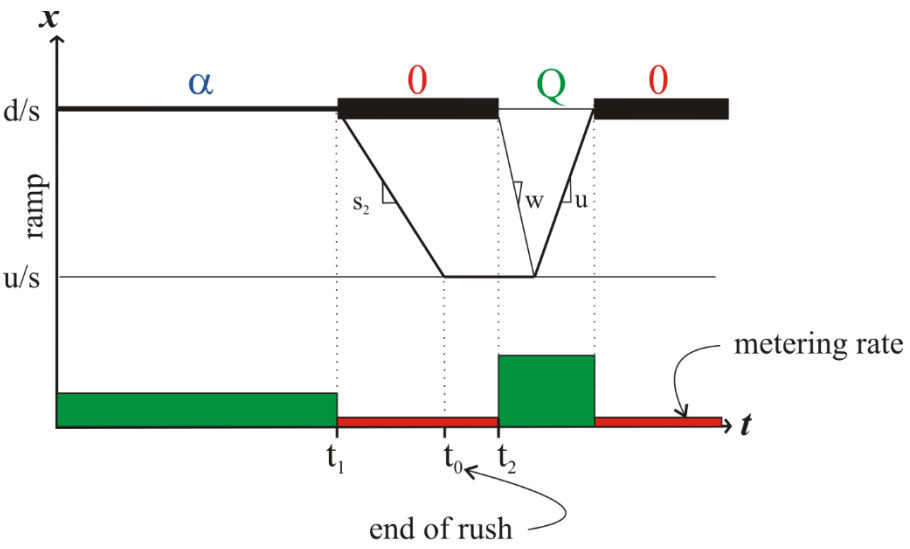

Figure 16: Optimal metering and ramp queues in Zone B

The ramp meters are shut down until the forward wave on the freeway, (due to the demand drop) arrives at the ramp and then the metering rate jumps to capacity flow. Unfortunately, an equation for the forward moving wave on the freeway is difficult to derive, to analytically calculate t_2. Once the wave *w* from t_2 hits the back of the queue on the ramp and the last vehicle leaves the ramp, the metering rate jumps to "capacity" flow.

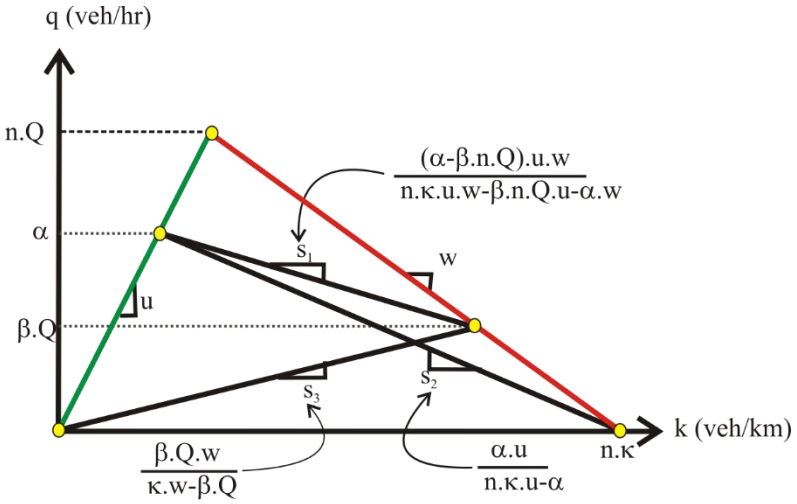

Figure 17: Shock speeds on the ramps during optimal control

Insight: *No metering until just before the demand drop, ramp closure until the forward shock reaches the ramp, then jump to capacity flow.* Such two-regime control is popularly known as *bang-bang* control [104].

Zone C

With Zone C located downstream of x_0 (see eqn. (14a) in Laval and Leclercq [106]) Zone 3 has transient congestion. As soon as the freeway reaches capacity, the optimal control rate becomes βQ, (see optimality conditions in section 4.1) resulting in a shock s_1 propagating backwards on the ramp (see Figure 18). Using equation (14b) in Laval and Leclercq [106] that determines the time when the freeway gets congested ($\triangleq t_5$) and equation for s_1 from Figure 17, equation for t_3 becomes:

$$t_3 = \frac{1}{\beta u}\ln\left(\frac{1}{1-(n\beta Q/\alpha)}\right) + \frac{l_r(n\kappa uw - \beta nQu - \alpha w)}{(\alpha - \beta nQ)uw} \qquad (56)$$

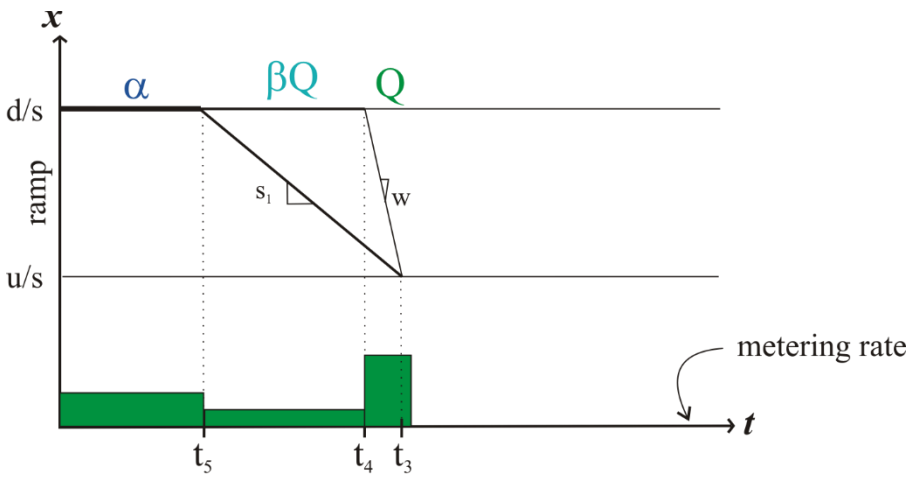

Figure 18: Optimal metering and ramp queues in Zone C

Then, two scenarios are possible depending on when the ramp empties:

Scenario 1: Ramp empties before s_1 reaches upstream

When t_0 falls between t_5 and t_3, (see Figure 19) a forward moving shock s_3 (see Figure 17) is generated at t_0 that intersects s_2 at:

$$t_1 = \frac{l_r + t_5 s_1 + t_0 u}{s_1 + u} \qquad (57)$$

and reaches freeway at:

$$t_2 = t_1 + \frac{s_1}{s_3}(t_1 - t_5) \qquad (58)$$

at which time the metering rate jumps to capacity flow. This is a simple scenario where the queue constraint is inconsequential.

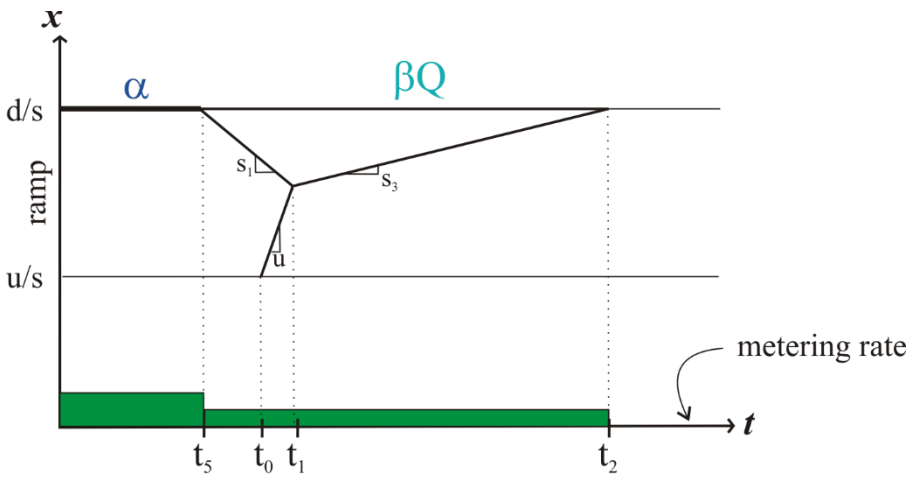

Figure 19: Optimal metering and ramp queues in Zone C – Scenario 1

Scenario 2: Shock s_1 reaches upstream and acceleration wave does not reach downstream

Following the shock s_1, to comply with the queue constraint, the metering rate jumps to capacity flow so the dissipation wave at t_4 reaches upstream just in time when s_1 reaches it. Thus, the equation for t_4 becomes:

$$t_4 = t_3 - \frac{l_r}{w} \qquad (59)$$

Since capacity flow decreases freeway efficiency, the metering rate switches back to *0*. At time t_0, when the demand drops, the queue stays in the middle of the ramp and Metering rate jumps to capacity when forward moving wave on the freeway reaches the ramp. For this scenario, the initiation and termination of the metering patterns are difficult to calculate analytically.

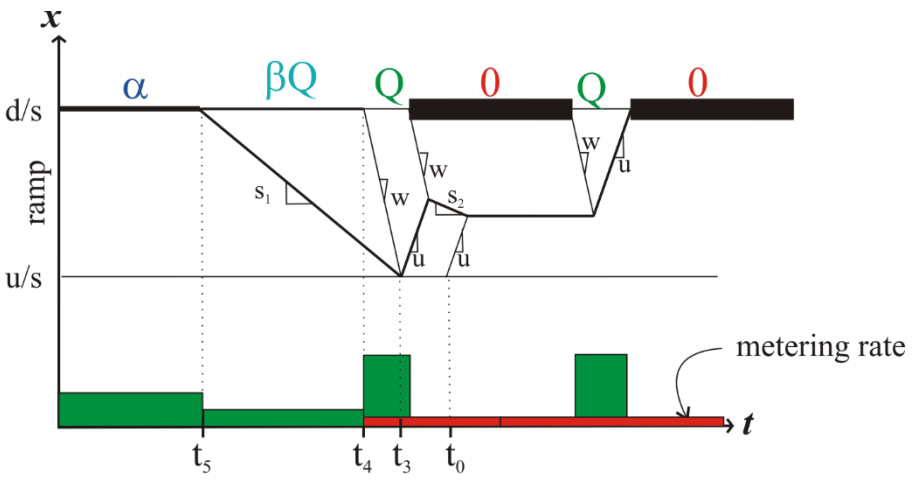

Figure 20: Optimal metering and ramp queues in Zone C – Scenario 2

Insight: *No metering until the freeway gets congested, meter at βQ to maintain capacity on the freeway and then turn off the meter to avoid violating queue constraint, turn off the ramp just before the demand drop and continue ramp closure until the forward shock reaches the ramp, then jump to capacity flow.*

Zone D

The optimal control pattern for Zone D during peak period loading is similar to that in Zone C (see Figure 21). However, unlike pattern in Zone C, the metering rate has βQ at the end that arises due to the ramp shutdown just before the demand drops.

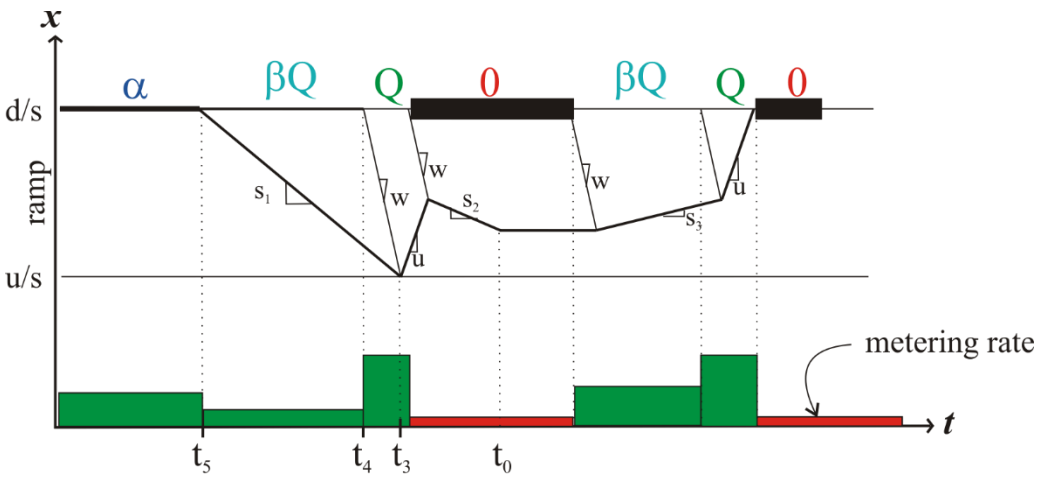

Figure 21: Optimal metering and ramp queues in Zone D

After the shockwave on the freeway reaches from downstream due to the ramp closure, the metering rate becomes $\beta\ Q$ and sustains until the forward moving disturbance reaches the ramp location. At that point, the metering goes up to capacity flow until the ramps are empty.

Insight: *No metering until the freeway gets congested, meter at βQ to maintain capacity on the freeway and then turn off the meter to avoid violating queue constraint, turn off the ramp just before the demand drop, jump to $\beta Q$when the freeway reaches capacity, then jump to capacity flow when the forward moving wave on the freeway reaches the ramp.*

Note that similar zones are observed with other values of a and b as long as there is freeway congestion, as shown in figures 22 and 23.

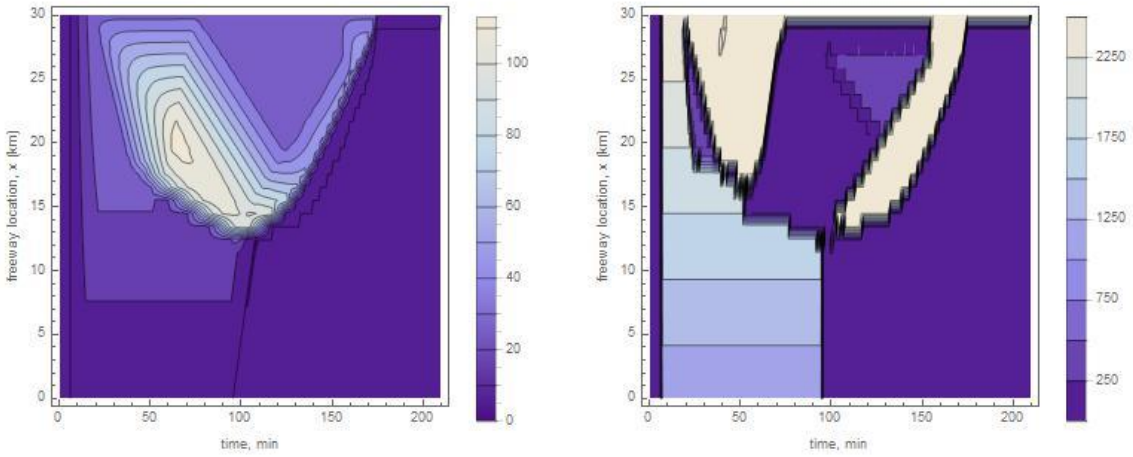

Figure 22: Relationship between optimal control strategy and state for increasing a and constant b

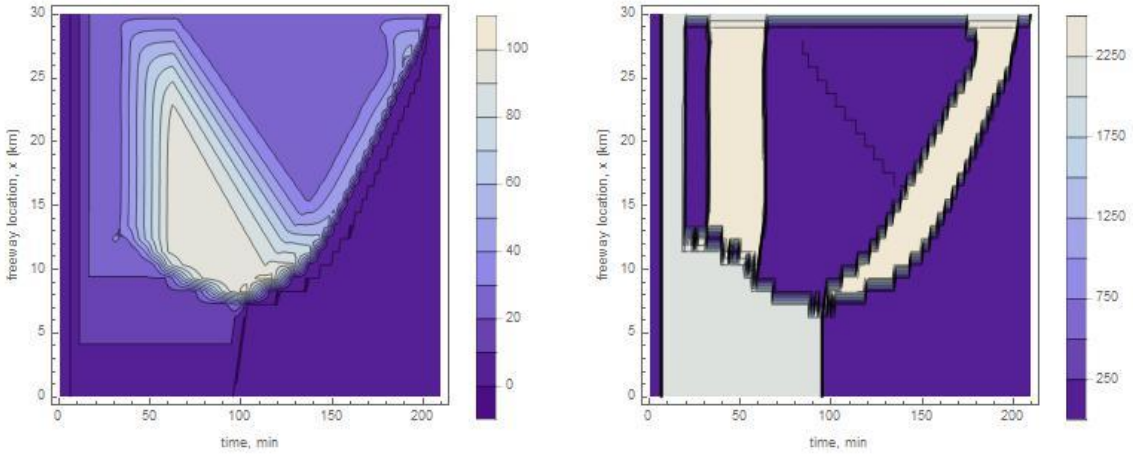

Figure 23: Relationship between optimal control strategy and state for increasing a and decreasing b

One of the features of the bang-bang control is that it is robustness to stochastic variations in the traffic demand as long as the disturbance does not reach the freeway during the

capacity flow period. We found that this phenomenon is true for ramp metering control with a 2-step time varying demand function.

6.2. Dimensionless Formulation

Problems with multiple variables can be efficiently evaluated using the *Dimensional Analysis* [126]. This technique involves reducing the number of variables by creating dimensionless combinations of the original variables using the Buckingham π theorem [127]. The theorem loosely states that if there is a physically meaningful problem with *n* physical variables, and *r* is the rank of the dimensional matrix, then, the problem can be reformulated with *n-r* dimensionless variables created from the original *n* variables that may not be unique and "physically" meaningful.

For the current problem: *time* and *space* are the basic variables and *inflow*, *outflow*, *freeway length, ramp length, ramp spacing, peak period duration, FD parameters*, etc. are the physical variables. We use this technique to derive universal shape of two parameters of the optimal control.

Example 1: Equation (14a) in Laval and Leclercq [106] defines the location of congestion initialization:

$$x_0 = \frac{1}{\beta}\ln\left(\frac{1}{1-(n\beta Q/\alpha)}\right) \qquad (60)$$

This equation is applicable to the current problem since control is implemented only at the onset of capacity conditions. Much more information can be extracted from *(60)* when rescaled with dimensionless variables:

$$\tilde{x}_0 = \frac{x_0}{l_f} \qquad b = \beta l_f \qquad a = \frac{\alpha l_f}{nQ} \qquad (61)$$

The $\tilde{x}_0$ represents the normalized $x_0$which indicates freeway congestion when its value is less than 1. b is the total outflow and a is the normalized total inflow. In the new coordinates, the equation *(62)* becomes:

$$\tilde{x}_0 = \frac{1}{b}\ln(\frac{1}{1-(b/a)}) \qquad (62)$$

Figure 24 shows the universal share of *(62),* which indicates that the congestion contours ($\tilde{x}_0$ <1) are nonlinear and cross y-axis at 1.

Insight*: The benefit of this transformation is that one only needs to keep track of a and b, in this case the system level aggregate variable values to calculate $\tilde{x}_0$ using (62). Then using Figure 24, one can determine if, where, and when the system will get congested. Note that this is true as long as the conditions for (63) hold true.*

Example 2: Similarly, the magnitude of shock s_1 is critical to the duration the ramp meter perform without queue override. We transfer the equation for s_1 using the following variables:

$$\tilde{s}_1 = \frac{s_1}{u} \qquad a = \frac{\alpha l_f}{nQ} \qquad b = \beta \qquad c = \frac{u}{w} \qquad (64)$$

The transformation means that the shock speed normalized with free-flow speed; lower $\tilde{s}_1$is desirable. With the transferred variables, the equation becomes:

$$\tilde{s}_1 = \frac{(a-b)}{(1+c-bc-a)} \qquad (65)$$

Figure 25 shows the universal shape of s_1 for different *c* values in *(66)* which indicates that in addition to *a* and *b,* the *u/w* ratio is critical and lower ratios give higher shock speeds that will result in frequent queue overrides. This is important in the era of vehicle-to-vehicle communication that could increase the wave speed.

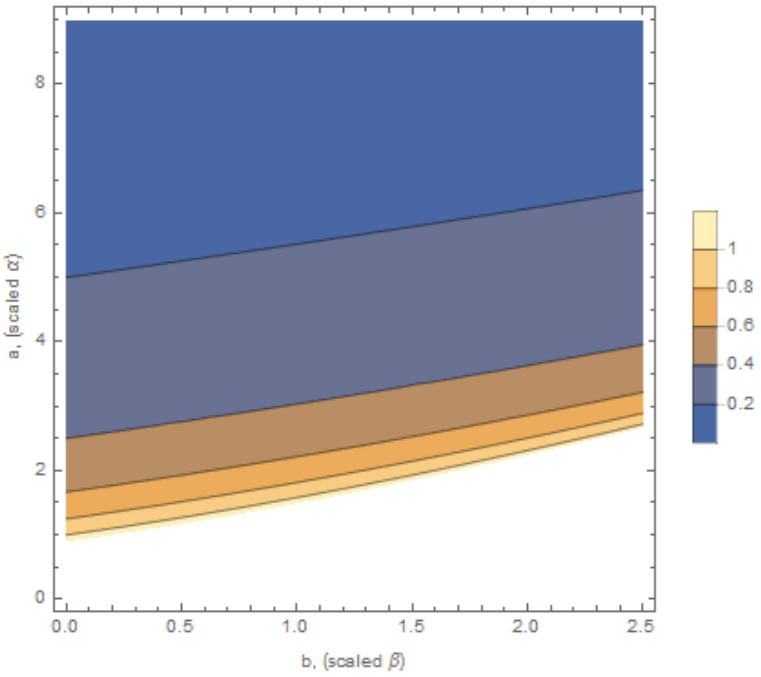

Figure 24: x_0 in rescaled dimensionless variables

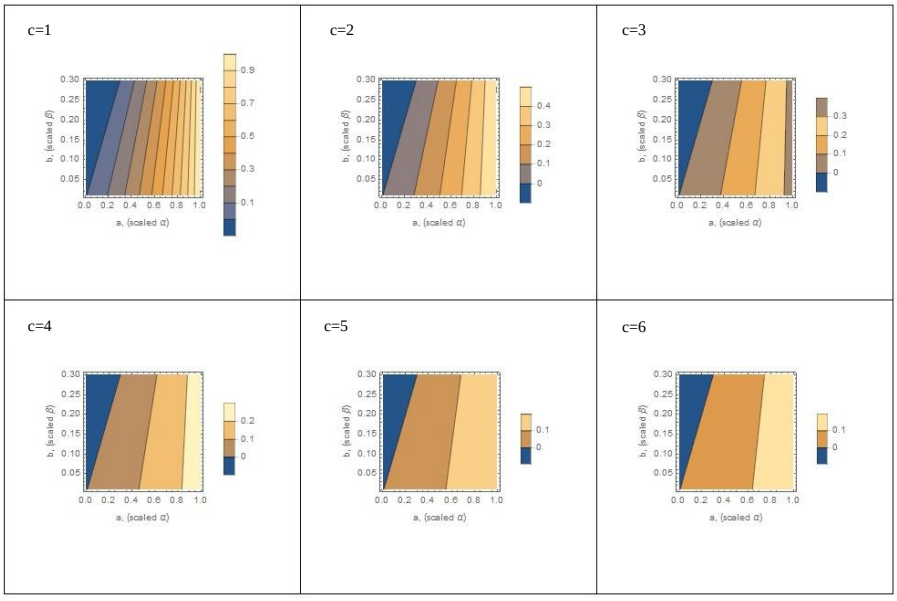

Figure 25: s_1 in rescaled variables

6.3. Merge Model Impacts

In this section, we evaluate the effects of merge model on the optimality conditions and derive bounds for the validity of the above findings.

The Newell-Daganzo merge model [61] (shown in Figure 26b) that was empirically proven by Anh and Cassidy [69] states that when the competing streams are congested, the available downstream capacity is shared at a fixed ratio independent of the (congested) flow in the competing approaches.

Assume a merge defined by the parameters shown in Figure 26a. the merge model in Figure 26b indicates that as long as the demand on the ramp, $\alpha\delta$, is within the green zone, the ramp priority is valid:

$$\alpha\delta \leq q_2 \triangleq \frac{\mu_3}{1+p} \qquad (67)$$

which is the upper bound for ramp demand.

The optimality conditions at the first queue spillback is to meter at capacity, Q. However, under such conditions, *(67)* indicates that the queue will form downstream of the meter. Subsequently, the tamp condition will deteriorate and queue spillback to the arterials becomes inevitable unless the ramp flow drops below the upper bound defined by *(67)*. One can find the bounds for b that ensures congestion using *(62)*:

$$0 < b < a + powerlog[\frac{-a}{e^a}] \qquad (68)$$

Note that *(67)* and *(68)* ensure that corridor is congested and the results obtained in the earlier section with ramp priority are valid.

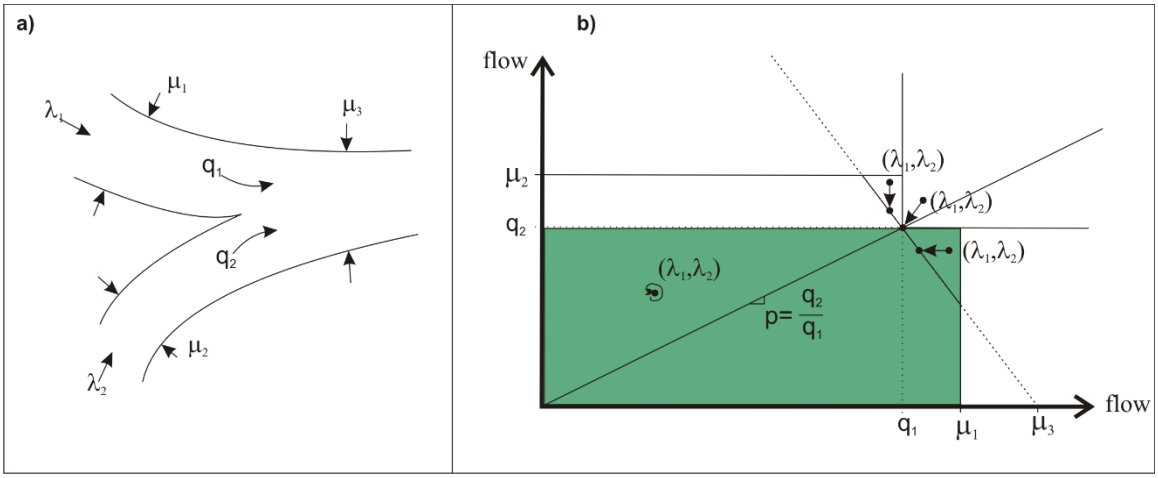

Figure 26: Region satisfying ramp priority on the merge model diagram a) parameters b) merge model

CHAPTER 7 HYBRID ISOLATED CONTROLLER

The first section of this chapter presents the philosophy of a new isolated ramp control developed based on the combined feedback and feedforward control. The next section presents an application of the hybrid algorithm that outperforms existing methods.

7.1. Philosophy

The optimality conditions in section 4.1 indicate that: a) metering should not be initiated until the freeway reaches capacity, b) the metering rate during capacity conditions is βQ, and c) the desired target state is k_c.

The literature review indicates that a) and b) can be efficiently achieved using a feedforward type control such as demand-capacity method. Moreover, c) can be achieved efficiently with a feedback type control such as ALINEA. Therefore, we devise a hybrid control from the feed-forward [16] and feedback [5] controllers (see equations *(1)*and*(2)*).

In the area of control theory, the demand-capacity type of controls (equation *(1)*) are classified as the feedforward disturbance rejection scheme. The strength of these strategies lies in their ability to quickly respond to any under-capacity conditions. However, they are known to be sensitive to unmeasured disturbances, i.e. moving bottlenecks, shockwave and change metering rates abruptly. Also, its control changes significantly and drastically only after the congestion sets in and the ramp metering rate is set to the minimum threshold that may sometimes lead to unnecessary under-loading of the freeway.

The popular feedback control, ALINEA (equation *(2)*), is an *I*-controller that corrects for the cumulative error, or deviation from the desired state. These controllers do

not need knowledge of the traffic flow models, but are robust, and provide a smooth transition during congestion. However, their reactive nature results in lag in responding to large disturbances and takes several iterations to restore back to the desired state.

The idea is to incorporate feed-forward control that has the traffic flow knowledge for quick response during free-flow conditions and use the feedback control for error correction during congestion. These types of methods are lately being used in medicine, aerospace, chemical processes etc. [128]–[132].

The block diagram of the hybrid control shown in Figure 27 indicates that the feedforward uses the upstream flow (q_{in}) in the flow conservation to calculate the metering rate. Feedback control uses the error ($O_{cr} - O_{out}$) and the rate in the previous time step to calculate the metering rate. The decision manager checks the traffic conditions and implements the former during free-flow conditions and the latter during congested conditions.

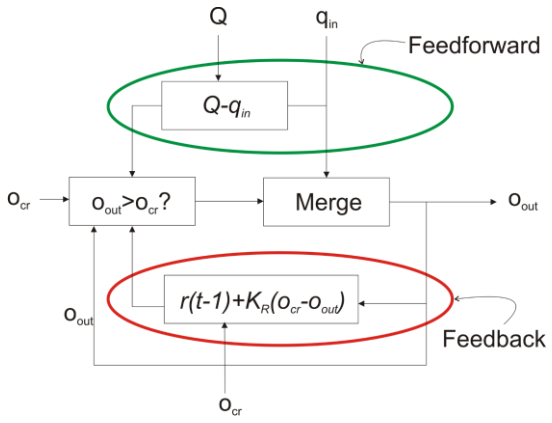

Figure 27: Block diagram of the hybrid isolated control

The hybrid control in Figure 27 is mathematically represents as follows:

$$r(t) = \begin{cases} Q - q_{in}(t-1), & O_{out} < O_{cr} \\ r(t-1) + K_R(O_{cr} - O_{out}(t-1)), & otherwise \end{cases} \quad (69)$$

Note that the hybrid control presented in this sections preserves the strengths of feed-forward and feed-back controllers and overcomes their limitations because the hybrid control responds quickly during free-flow and allows for a smooth transition during congestion.

7.2. Application

This section presents the results of the application of the hybrid control on a 20 kilometer long 1-lane freeway corridor with the ramps spaced 1 km apart. We load the system for 1 hour with a demand of 2000 veh/hr at each ramp and use an exit proportion of 25%. The system is simulated with a triangular FD (u = 100 kph, w = 20 kph, and κ = 150 veh/km), until it is empty. Note that we use the merge model from Laval and Leclercq [106] in this analysis.

The cumulative departures at the ramp meter shown in Figure 28 indicate that "no metering" pushes more vehicles early on but flattens out later due to the spillback from the downstream of the meter (merge model governs the flows entering the freeway).

Both ALINEA and hybrid control perform almost identically until the effects of demand drop are felt. ALINEA responds slowly, while the Hybrid control pushes more vehicles into the freeway.

The cumulative departures from the system in Figure 29 indicate that "no metering" resulted in overloading of the freeway that reduced the system outflow. Initially, ALINEA and the Hybrid control maintain similar freeway densities, but when the demand dropped, the Hybrid control responded quickly and the system outflows increased. The delay benefits of the Hybrid control over the ALINEA are almost 8%.

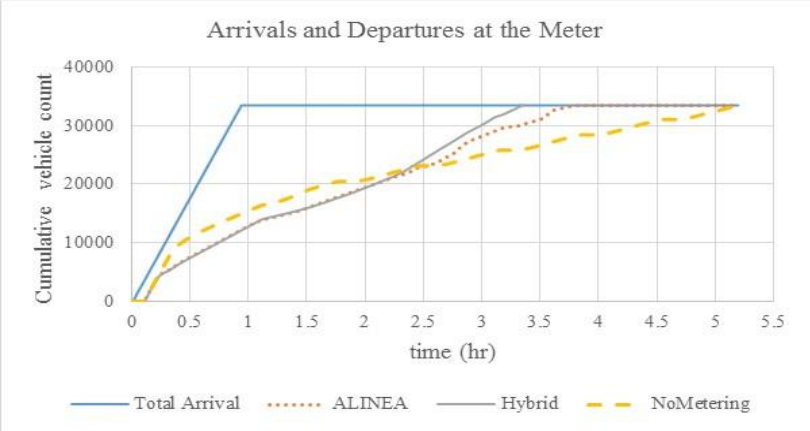

Figure 28: Cumulative arrivals and departures at the ramp meter; No metering, ALINEA, and Hybrid control

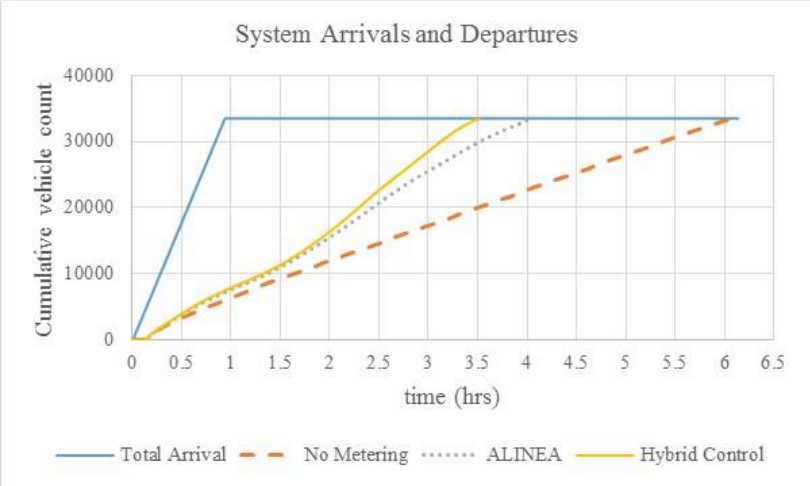

Figure 29: Time-space plot of the discrete ramps. a) no metering b) ALINEA c) Hybrid Control

The freeway congestion dynamics presented in Figure 30 confirm that the ALINEA and the Hybrid control perform almost identically initially, but the latter outperforms the former by reducing freeway congestion and emptying the system early.

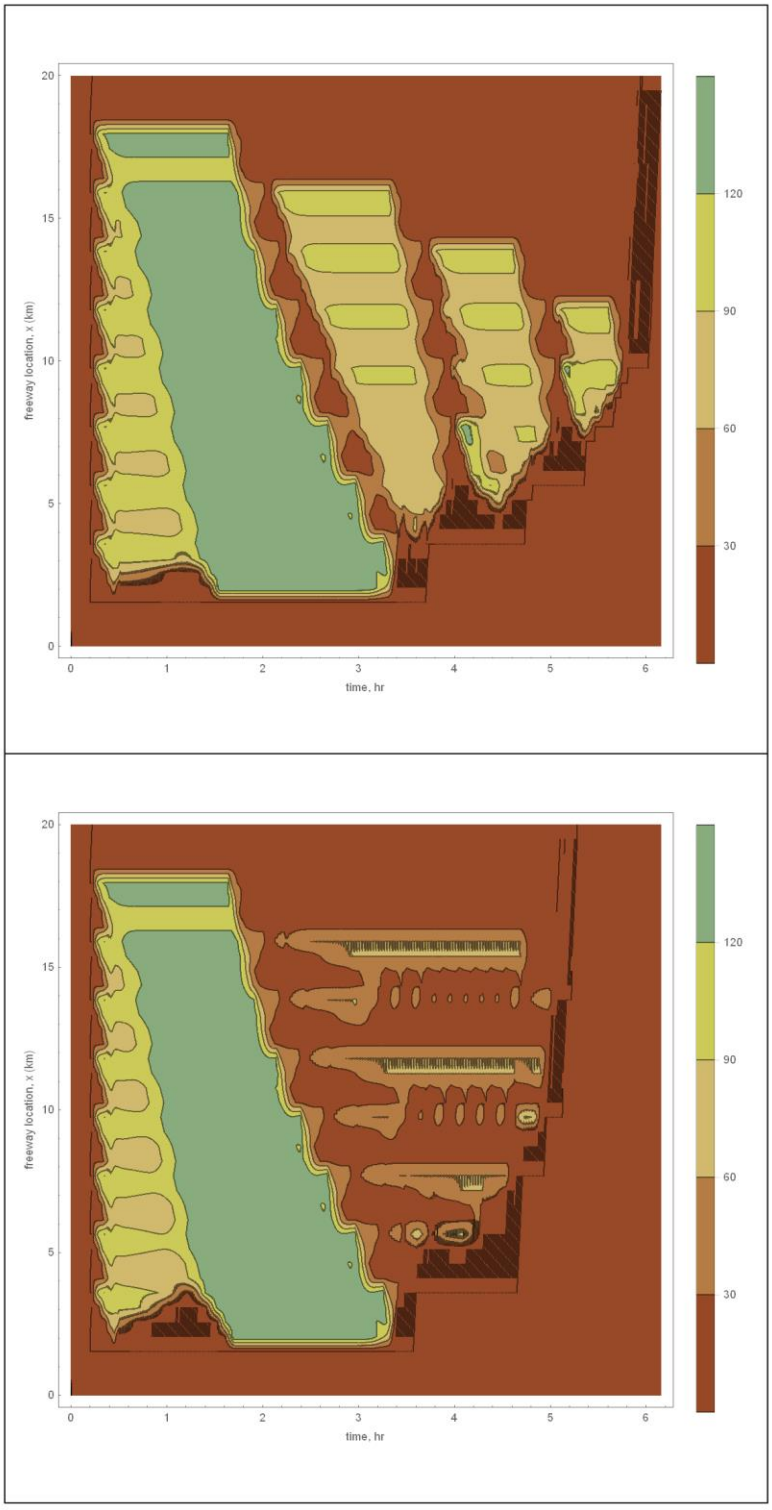

Figure 30: Time-space density plot a) ALINEA b) Hybrid Control

Since hybrid control is quick to respond, it is better equipped to efficiently handle stochasticity in traffic demand. However, one should note that compared to the ALINEA, hybrid control needs more infrastructure to implement in the field; switching manager and extra detector for upstream inflow.

Based on the lessons learned in chapter 4, this chapter presented a hybrid isolated controller that outperformed existing methods. The next chapter will present a discussion of the findings, contributions, and extensions to the current work.

CHAPTER 8 DISCUSSION AND EXTENSIONS

This dissertation studied the optimality conditions for constrained and unconstrained on-ramp queues to unveil interesting qualitative aspects of the optimal control trajectory and its correlation with state trajectories on the freeway and on-ramps.

This research, for the first time, explicitly modeled the special dynamics at on-ramps as part of the optimal control problem. This was critical for explaining important relationships between the control and traffic dynamics that until now were not possible since ramps were modeled as vertical queues in the literature. However, one of the challenges with modeling ramps is the increase in the number of variables and constraints for ramps, but the cost is justified due to insights gained.

We find that the freeway operated at capacity is the system-wide optimality condition, which can be attained with local metering in the case of unconstrained queues. This complies with the commonly accepted "heuristic" theory that the ramp meters should aim to maintain freeway at capacity. Similarly, when there is queue constraint, we observe that the queue should be flushed, as it is common practice today. However, at the end of the peak period, we have shown that the optimal strategy is to close the ramp to alleviate freeway congestion and turn-off the metering when the forward moving wave on the freeway reaches the ramp location. This bang-bang-type control opens the door for a new set of algorithms that are robust, optimal and simple to implement. Moreover, this study also identified four zones that have distinct metering patterns, which are correlated with the shock evolution on the freeway and ramps.

The error-free EG scheme presented in section 4.1 is an extension of the variational solution for KW model with source term for the special case of homogenous

corridor with triangular fundamental diagram. It is shown that the proposed scheme provides an improved numerical solution compared to the classical Godunov scheme.

The hybrid control presented in chapter 6 is significant to the state-of-the-art practice. It combines the well-established feedback and feedforward controllers, preserving their strengths and overcoming their limitations. With its quick response, it is conjectured that this controller has the potential to handle planned and unplanned events effectively. Note that the idea of hybrid control proposed in this research is supported by the classical control theory that recommends supplementing feedback control with a feedforward control (that has process information) to obtain better performance [133].

One of the limitations of this study is that it assumes ramp priority to circumvent the nonlinearity of the N-D merge model and obtain global optimal results with the LP formulation. However, as discussed in section 6.3, this leads to the results being valid only for a set of conditions.

8.1. Contributions

The major contributions of this research are as follows:

1) It provides optimality conditions for the case of unlimited ramp queues.
2) It presents an error-free solution to the Kinematic Wave model with a source term.
3) It provides an insight into why simulation-based optimization may not be a suitable framework for optimal control problems in traffic flow.
4) It identifies regions in the state surface that have a distinct control patterns for the case of queue constraint and unveils the correlation between the

spatiotemporal control trajectory and the shockwave evolution on the freeway and ramps.

5) It proposes a hybrid isolated control that outperforms existing methods.

8.2. Extensions

This research can be extended in several ways:

Solve Continuum Formulation: The continuum formulation in section 4.2 can be relaxed to provide priority to the ramp flows and solved with variational methods.

Incorporate Merge Model: The merge model may be incorporated by reposing the model to fit the LP or quadratic programming framework.

Discrete ramps: The dynamics of the discrete ramps may be different and developing real-time controllers will need an understanding of the microscopic relationships.

Stochastic demands: This study only investigated the space-varying demand patterns. However, time-varying demands whose influence can be felt downstream of the meter will be more realistic and interesting to study.

Equity: Literature is conclusive to show that optimal control is not equitable [134], but efficient ways to balance equity and optimality is still an open question.

Efficient queue override mechanism: The convenience with optimization formulations is that spillback can be prevented with a simple constraint equation. In practice, this is extremely difficult and need to be investigated.

Coordinated and Integrated control: A study with discrete-ramp modeling with stochastic demands will exacerbate the ill effects of capacity drop and require the need for efficient management of storage. This would open doors to a new set of methods that coordinates across ramps and integrates across devices to obtain system optimality.

Note that these results are an open loop optimal, which is the best solution possible because the traffic demands are a priori. However, this control cannot be implemented in the field as is. This analysis was done to gain a qualitative insight into optimality conditions. However, for systems where optimality conditions are possible to identify, one could use a model predictive control-like frameworks to obtain optimality in real-time.

APPENDIX A. OPTIMALITY CONDITIONS WITH ALTERNATIVE FORMULATION

Let the density on a long freeway-ramp system are given by $k_f(t,x)$and $k_r(t,x,y)$, where the subscript *f* denotes freeway, *r* denotes the ramp, and *x* and *y* denote length along freeway and ramp respectively. If the length of the freeway, ramp and the optimization duration are L_x, L_y, and *T* respectively, the TTS is given by,

$$TTS = \int_0^T \int_0^{L_x} k_f(t,x)\,dx\,dt + \int_0^T \int_0^{L_x} \int_0^{L_y} k_r(t,x,y)\,dx\,dy\,dt \qquad (70)$$

Since the $k \triangleq -\frac{\partial N_f}{\partial x}\ or\ -\frac{\partial N_r}{\partial y}$, where N_i is the vehicle count, *(70)* can be re written as

$$-TTS = \int_0^T N_f(t,L_x)\,dt + \int_0^T N_r\big(t,L_x,L_y\big)\,dt \qquad (71)$$

where $N_f(t)$ and $N_r(t)$ are the total number of vehicles queued on the freeway and ramps. This implies that:

$$minimize\ TTS \Longrightarrow maximize \int_0^T N(t)\,dt \qquad (72)$$

Hence *(72)* is equivalent to *(28)*.

REFERENCES

[1] M. Papageorgiou and A. Kotsialos, "Freeway ramp metering: An overview," *IEEE Trans. Intell. Transp. Syst.*, vol. 3, no. 4, pp. 271–281, 2002.

[2] J. a Wattleworth, "Peak-period analysis and control of a freeway system," *Highw. Res. Rec.*, no. 89, pp. 1–25, 1965.

[3] J. B. Yuan L.S. and Kreer, "An optimal control algorithm for ramp metering of urban freeways," in *6th Annual Allerton Conference on Circuit and System Theory*, 1968.

[4] J. J. Wang and A. D. May, "Computer model for optimal freeway on-ramp control," *Highw. Res. Rec.*, vol. 469, no. 469, pp. 16–25, 1973.

[5] M. Papageorgiou, H. Hadj-Salem, and J. M. Blosseville, "Alinea. A local feedback control law for on-ramp metering," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1320, pp. 58–64, 1991.

[6] E. Smaragdis and M. Papageorgiou, "Series of New Local Ramp Metering Strategies," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1856, pp. 74–86, 2003.

[7] Y. J. Stephanedes and K. K. Chang, "Optimal control of freeway corridors," *J. Transp. Eng.*, vol. 119, no. 4, pp. 504–514, 1993.

[8] H. M. Zhang, S. G. Ritchie, and W. W. Recker, "some General results on the optimal ramp control problem," *Transp. Res. Part C*, vol. 4, no. 2, pp. 51–69, 1996.

[9] A. Kotsialos, M. Papageorgiou, and F. Middelham, "Optimal Coordinated Ramp Metering with Advanced Motorway Optimal Control," *Transp. Res. Rec.*, vol. 1748, no. 1, pp. 55–65, Jan. 2001.

[10] Y. J. Stephanedes, "Implementation of on-line Zone Control Strategies for optimal ramp metering in the Minneapolis Ring Road," in *Road Traffic Monitoring and Control, 1994., Seventh International Conference on*, 1994, pp. 181–184.

[11] G. Paesani, J. Kerr, P. Perovich, and F. E. Khosravi, “System wide adaptive ramp metering (SWARM),” *Merging Transp. Commun. Revolutions. Abstr. ITS Am. Seventh Annu. Meet. Expo.*, 1997.

[12] I. Papamichail, M. Papageorgiou, V. Vong, and J. Gaffney, “Heuristic Ramp-Metering Coordination Strategy Implemented at Monash Freeway, Australia,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2178, pp. 10–20, 2010.

[13] L. Jacobson, K. Henry, and O. Mehyar, “Real-time metering algorithm for centralized control,” *Transp. Res. Rec.*, no. 1732, pp. 20–32, 1989.

[14] L. Chu, H. X. Liu, W. Recker, and H. M. Zhang, “Performance Evaluation of Adaptive Ramp-Metering Algorithms Using Microscopic Traffic Simulation Model,” *J. Transp. Eng.*, vol. 130, no. 3, pp. 330–338, 2004.

[15] M. Papageorgiou and I. Papamichail, “Handbook of ramp metering,” *Doc. Eur. Ramp Metering Proj. (…*, no. 507645, 2007.

[16] D. P. Masher, D. W. Ross, P. J. Wong, P. L. Tuan, H. M. Zeidler, and S. Peracek, “Guidelines for design and operating of ramp control systems.,” Menlo Park, California, 1975.

[17] M. Papageorgiou, H. Hadj-Salem, and F. Middelham, “ALINEA Local Ramp Metering: Summary of Field Results,” *Transp. Res. Rec.*, vol. 1603, no. 1, pp. 90–98, Jan. 1997.

[18] C. Taylor, D. Meldrum, and L. Jacobson, “Fuzzy ramp metering: design overview and simulation results,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1634, no. -1, pp. 10–18, 1998.

[19] G. Paesani, J. Kerr, P. Perovich, and F. E. Khosravi, “System wide adaptive ramp metering (SWARM),” in *Merging the Transportation and Communications Revolutions. Abstracts for ITS America Seventh Annual Meeting and Exposition*, 1997.

[20] I. Papamichail and M. Papageorgiou, “Traffic-Responsive Linked Ramp-Metering Control,” *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 1, pp. 111–121, 2008.

[21] E. Smaragdis, M. Papageorgiou, and E. Kosmatopoulos, “A flow-maximizing

adaptive local ramp metering strategy," *Transp. Res. Part B*, vol. 38, no. 3, pp. 251–270, 2004.

[22] B. Feng, J. Hourdos, and P. Michalopoulos, "Improving Minnesota's Stratified Ramp Control Strategy," *Transp. Res. Rec.*, vol. 1959, no. 1, pp. 77–83, Jan. 2006.

[23] H. M. Zhang, S. G. Ritchie, and R. Jayakrishnan, "Coordinated traffic-responsive ramp control via nonlinear state feedback," *Transp. Res. Part C*, vol. 9, no. 5, pp. 337–352, 2001.

[24] A. Messmer and M. Papageorgiou, "Motorway Network Control via Nonlinear optimization," *Int. Trans. Oper. Res.*, vol. 2, no. 2, pp. 187–203, 1995.

[25] a. Kotsialos and M. Papageorgiou, "Nonlinear optimal control applied to coordinated ramp metering," *IEEE Trans. Control Syst. Technol.*, vol. 12, no. 6, pp. 920–933, 2004.

[26] I. Papamichail, A. Kotsialos, I. Margonis, and M. Papageorgiou, "Coordinated ramp metering for freeway networks – A model-predictive hierarchical control approach," *Transp. Res. Part C Emerg. Technol.*, vol. 18, no. 3, pp. 311–331, 2008.

[27] a Alessandri, "Optimal control of freeways via speed signalling and ramp metering," *Control Eng. Pract.*, vol. 6, no. 6, pp. 771–780, Jun. 1998.

[28] a Kotsialos, "Coordinated and integrated control of motorway networks via non-linear optimal control," *Transp. Res. Part C Emerg. Technol.*, vol. 10, no. 1, pp. 65–84, Feb. 2002.

[29] C.-H. Wei, "Analysis of artificial neural network models for freeway ramp metering control," *Artif. Intell. Eng.*, vol. 15, no. 3, pp. 241–252, 2001.

[30] J. A. Wattleworth, "Peak-period analysis and control of a freeway system," 1965.

[31] R. L. Pretty, "Control of a Freeway System by Means of Ramp Metering and Information Signs," *Highw. Res. Rec.*, vol. 388, 1972.

[32] J. A. Wattleworth and D. S. Berry, "Peak-period Control of a Freeway System - Some Theoretical Investigations," *Highw. Res. Rec.*, vol. 89, pp. 1–25, 1965.

[33] H. M. Zhang and S. Ritchie, “Freeway ramp metering using artificial neural networks,” *Transp. Res. Part C Emerg. …*, vol. 5, no. 5, pp. 273–286, 1997.

[34] K. Bogenberger, S. Vukanovic, and H. Keller, “ACCEZZ—Adaptive Fuzzy Algorithms for Traffic Responsive and Coordinated Ramp Metering,” in *Journal of Transportation Engineering*, 2002, vol. 49, no. 0, pp. 744–753.

[35] S. C. F. Iida Y., Hasegawa T., Asakura Y., “A formulation of on-ramp traffic control system with route guidance for urban expressway,” in *IFAC/IFFIP/IFORS/-Sixth International Conference on Control in Transportation Systems*, 1989, pp. 229–236.

[36] H. Yang, S. Yagar, Y. Iida, and Y. Asakura, “An algorithm for the inflow control problem on urban freeway networks with user-optimal flows,” *Transp. Res. Part B …*, vol. 28, no. 2, pp. 123–139, 1994.

[37] J. C. M. Baños and M. Papageorgiou, “A Linear Programming Approach to Large-Scale Linear Optimal Control Problems,” *IEEE Trans.\ Aut.\ Control*, vol. 40, no. 5, pp. 971–977, 1995.

[38] M. Papageorgiou, “A new approach to time-of-day control based on a dynamic freeway traffic model,” *Transp. Res. Part B Methodol.*, vol. 14, no. 4, pp. 349–360, 1980.

[39] M. Papageorgiou, E. Kosmatopoulos, I. Papamichail, and Y. Wang, “A Misapplication of the Local Ramp Metering Strategy ALINEA,” *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 2, pp. 360–364, 2008.

[40] Y. Wang, E. B. Kosmatopoulos, M. Papageorgiou, and I. Papamichail, “Local Ramp Metering in the Presence of a Distant Downstream Bottleneck: Theoretical Analysis and Simulation Study,” *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2024–2039, 2014.

[41] D. Lovell and C. F. Daganzo, “Access control on networks with unique origin–destination paths,” *Transp. Res. Part B Methodol.*, vol. 34, no. 3, pp. 185–202, Apr. 2000.

[42] G. Gomes and R. Horowitz, “Optimal freeway ramp metering using the

asymmetric cell transmission model," *Transp. Res. Part C Emerg. …*, vol. vol, pp. 14no4pp244–262, 2006.

[43] L. Isaksen and H. Payne, "Suboptimal control of linear systems by augmentation with application to freeway traffic regulation," *IEEE Trans. Automat. Contr.*, vol. AC-18, no. 3, pp. 210–219, 1973.

[44] N. B. Goldstein and K. S. P. Kumar, "A decentralized control strategy for freeway regulation," *Transp. Res. Part B Methodol.*, vol. 16, no. 4, pp. 279–290, 1982.

[45] O. J. Chen, A. F. Hotz, and M. E. Ben-Akiva, "Development and evaluation of a dynamic ramp metering control model," in *Proceedings of the 8th IFAC/IFIP/IFORS. Transportation Systems 1997 (3 vol.), 16-18 June 1997*, 1997, vol. vol.3, pp. 1089–1095.

[46] C. Pasquale, I. Papamichail, C. Roncoli, S. Sacone, S. Siri, and M. Papageorgiou, "Two-class freeway traffic regulation to reduce congestion and emissions via nonlinear optimal control," *Transp. Res. Part C Emerg. Technol.*, vol. 55, pp. 85–99, 2015.

[47] R. C. Carlson, I. Papamichail, M. Papageorgiou, and A. Messmer, "Optimal mainstream traffic flow control of large-scale motorway networks," *Transp. Res. Part C Emerg. Technol.*, vol. 18, no. 2, pp. 193–212, 2010.

[48] C. Roncoli, M. Papageorgiou, and I. Papamichail, "Traffic flow optimisation in presence of vehicle automation and communication systems – Part II: Optimal control for multi-lane motorways," *Transp. Res. Part C Emerg. Technol.*, vol. 57, pp. 260–275, 2015.

[49] C. Diakaki and M. Papageorgiou, "Design and Simulation Test of Coordinated Ramp Metering Control (METALINE) for A10-West in Amsterdam," ATT-Project Eurocor V2017, 1994.

[50] M. Papageorgiou, B. Jean-marc, and H. Hadj-salem, "Macroscopic Modelling of traffic flow on the Boulevard Peripherique in Paris," *trafnsportation Res. PartB Methodol.*, vol. 236, no. ii, pp. 29–47, 1989.

[51] M. Papageorgiou, J.-M. Blosseville, and H. Haj-Salem, "Modelling and real-time control of traffic flow on the southern part of Boulevard Peripherique in Paris: Part II: Coordinated on-ramp metering," *Transp. Res. Part A Gen.*, vol. 24, no. 5, pp.

361–370, 1990.

[52] M. Papageorgiou, J.-M. Blosseville, and H. Haj-Salem, “Modelling and real-time control of traffic flow on the southern part of Boulevard Peripherique in Paris: Part II: Coordinated on-ramp metering,” *Transp. Res. Part A Gen.*, vol. 24, no. 5, pp. 361–370, 1990.

[53] H. M. Zhang and W. W. Recker, “On optimal freeway ramp control policies for congested traffic corridors,” *Transp. Res. Part B Methodol.*, vol. 33, no. 6, pp. 417–436, 1999.

[54] R. C. Carlson, I. Papamichail, M. Papageorgiou, and A. Messmer, “Optimal Motorway Traffic Flow Control Involving Variable Speed Limits and Ramp Metering,” *Transp. Sci.*, vol. 44, no. 2, pp. 238–253, 2010.

[55] M. Papageorgiou, “An integrated control approach for traffic corridors,” *Transp. Res. Part C Emerg. Technol.*, vol. 3, no. 1, pp. 19–30, 1995.

[56] J. H. Banks, “Effect of response limitations on traffic-responsive ramp metering,” *Transp. Res. Rec.*, no. 1394, 1993.

[57] L. E. Lipp, L. J. Corcoran, and G. A. Hickman, “Benefits of central computer control for Denver ramp-metering system,” *Transp. Res. Rec.*, no. 1320, 1991.

[58] Cambridge Systematics, “Twin Cities Ramp Meter Evaluation Final Report,” 2001.

[59] Pham H., “SWARM Study Final Report on W/B Foothill Freeway,” 2002.

[60] L. Newman, A. Dunnet, and G. . Meis, “Freeway ramp control - what it can and can not do,” *Traffic Eng.*, vol. 39, pp. 14–21, 1969.

[61] C. F. Daganzo, “The nature of freeway gridlock and how to prevent it,” *Proc. 13th Int. Symp. Transp. Traffic Theory*, pp. 629–646, 1996.

[62] B. D. Greenshields, J. Bibbins, W. Channing, and H. Miller, “A study of traffic capacity,” *Highway Research Board Proceedings*. pp. 448–477, 1935.

[63] M. Carey and M. Bowers, “A Review of Properties of Flow–Density Functions,” *Transp. Rev.*, vol. 32, no. 1, pp. 49–73, 2012.

[64] M. J. Cassidy, “Bivariate relations in nearly stationary highway traffic,” *Transp. Res. Part B*, vol. 32, no. 1, pp. 49–59, 1998.

[65] R. L. Bertini and M. J. Cassidy, “Some observed queue discharge features at a freeway bottleneck downstream of a merge,” *Transp. Res. Part A*, vol. 36, no. 8, pp. 683–697, 2002.

[66] J. C. Munoz and C. Daganzo, “Moving bottlenecks: a theory grounded on experimental observation,” in *Proceedings of the 15th International Symposium on Transportation and Traffic Theory (ISTTT15)*, 2002, pp. 441–462.

[67] M. J. Cassidy and J. R. Windover, “Methodology for assessimg dynamics of freeway traffic flow,” *Transp. Res. Rec. J. Transp. Res. Board*, no. 1484, pp. 73–79, 1995.

[68] K. R. Smilowitz, C. F. Daganzo, M. J. Cassidy, and R. L. Bertini, “Some observations of highway traffic in long queues,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1678, no. -1, pp. 225–233, 1999.

[69] M. J. Cassidy and S. Ahn, “Driver Turn-Taking Behavior in Congested Freeway Merges,” *Transp. Res. Rec.*, vol. 1934, pp. 140–147, 2005.

[70] A. Kotsialos and M. Papageorgiou, “The importance of traffic flow modeling for motorway traffic control,” *Networks Spat. Econ.*, pp. 179–203, 2001.

[71] L. Zhang and D. Levinson, “Optimal freeway ramp control without origin–destination information,” *Transp. Res. Part B Methodol.*, vol. 38, no. 10, pp. 869–887, 2004.

[72] M. J. Lighthill and G. B. Whiteham, “On kinematic waves. I. Flood movement in long rivers. II. A Theory of Traffic Flow on Long Crowded Roads,” *Proc. R. Soc. London, Ser A*, vol. 299, no. 1178, p. 281\,–\,345, 1955.

[73] P. I. Richards, “Shock Waves on the Highway,” *Oper. Res.*, vol. 4, no. 1, pp. 42–51, 1956.

[74] R. J. LeVeque, *Numerical Methods for Conservation Laws*, vol. 57. 1992.

[75] J. Lebacque, "Two-Phase Bounded-Acceleration Traffic Flow Model: Analytical Solutions and Applications," *Transportation Research Record*, vol. 1852, no. 1. pp. 220–230, 2003.

[76] S. K. Godunov, "A Difference Scheme for Numerical Solution of Discontinuous Solution of Hydrodynamic Equations," *Mat. Sb.*, vol. 89, no. 3, pp. 271–306, 1959.

[77] C. F. Daganzo, "The Cell Transmission Model. Part I: A Simple Dynamic Representation of Hiway Traffic," *Transp. Res. Part B Methodol.*, pp. 1055–1425, 1993.

[78] J.-P. Lebacque, "The Godunov scheme and what it means for first order traffic flow models," *Int. Symp. Transp. Traffic Theory*, pp. 647–677, 1996.

[79] R. J. LeVeque, "Finite Volume Methods for Hyperbolic Problems," *Cambridge Univ. Press*, vol. 54, p. 258, 2002.

[80] J. B. Edwards, "Speed adjustment of motorway commuter traffic to inclement weather," *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 2, no. 1, pp. 1–14, 1999.

[81] H. Rakha, M. Farzaneh, M. Arafeh, and E. Sterzin, "Inclement Weather Impacts on Freeway Traffic Stream Behavior," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2071, no. -1, pp. 8–18, Dec. 2008.

[82] J. a Laval and L. Leclercq, "A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic.," *Philos. Trans. A. Math. Phys. Eng. Sci.*, vol. 368, no. 1928, pp. 4519–41, Oct. 2010.

[83] D. Samba and B. Park, "Incorporating Weather Impacts in Traffic Estimation and Prediction Systems," *Proc. 18th ITS World Congr. Intell. Transp. Syst.*, pp. 1–22, 2011.

[84] K. Agyemang-Duah and F. Hall, "Some issues regarding the numerical value of freeway capacity," in *International Symposium on Highway Capacity*, 1991.

[85] M. J. Cassidy and R. L. Bertini, "Some traffic features at freeway bottlenecks,"

Transp. Res. Part B Methodol., vol. 33, no. 1, pp. 25–42, 1999.

[86] B. R. Chilukuri, J. a Laval, and D. Chen, "Some Traffic Features During On-ramp Queue Flush," in *Transportation Research Board 92nd Annual Meeting*, 2013, pp. 1–15.

[87] L. Leclercq, J. a. Laval, and N. Chiabaut, "Capacity drops at merges: An endogenous model," *Transp. Res. Part B Methodol.*, vol. 45 `, no. 9, pp. 1302–1313, 2011.

[88] L. Leclercq, V. L. Knoop, F. Marczak, and S. P. Hoogendoorn, "Capacity drops at merges: New analytical investigations," *Transp. Res. Part C*, 2015.

[89] Y. Lu, S. C. Wong, M. Zhang, C. W. Shu, and W. Chen, "Explicit construction of entropy solutions for the Lighthill-Whitham-Richards traffic flow model with a piecewise quadratic flow-density relationship," *Transp. Res. Part B Methodol.*, vol. 42, no. 4, pp. 355–372, 2008.

[90] Y. Lu, S. C. Wong, M. Zhang, and C.-W. Shu, "The Entropy Solutions for the Lighthill-Whitham-Richards Traffic Flow Model with a Discontinuous Flow-Density Relationship," *Transp. Sci.*, 2009.

[91] W.-L. Jin, Q.-J. Gan, and J.-P. Lebacque, "A kinematic wave theory of capacity drop," *Transp. Res. Part B*, p. 29, 2015.

[92] M. Hadiuzzaman and T. Qiu, "Cell transmission model based variable speed limit control for freeways," *Can. J. Civ. Eng.*, vol. 40, no. 1, pp. 46–56, 2013.

[93] L. Leclercq, J. Laval, and E. Chevallier, "The Lagrangian coordinate system and what it means for first order traffic flow models," *17th Int. Symp. Transp. Traffic Theory*, no. Section 2, pp. 1–13, 2007.

[94] G. Newell, "A simplified theory of kinematic waves in highway traffic, part I: general theory," *Transp. Res. Part B Methodol.*, vol. 27, no. 4, pp. 281–287, Aug. 1993.

[95] C. F. Daganzo, "A Variational Formulation of Kinematic Wave Theory: basic theory and complex boundary conditions," *Transp. Res. Part B*, vol. 39, no. 2, pp. 187–196, 2003.

[96] C. F. Daganzo, “On the variational theory of traffic flow: well-posedness, duality and applications,” *Networks Heterog. Media*, vol. 1, pp. 601–619, 2006.

[97] Y. Makigami, G. Newell, and R. Rothery, “Three-dimensional representation of traffic flow,” *Transp. Sci.*, vol. 5, no. 3, p. 302, 1971.

[98] C. F. Daganzo, “A variational formulation of kinematic waves: Solution methods,” *Transp. Res. Part B Methodol.*, vol. 39, pp. 934–950, 2005.

[99] Payne H., “Models of freeway tra_c and control,” *Math. Model. public Syst.*, vol. 1, no. 1, pp. 51–61, 1971.

[100] D. Helbing, “Gas-kinetic derivation of Navier-Stokes-like traffic equations,” *Phys. Rev. E*, vol. 53, no. 3, pp. 2366–2381, 1996.

[101] C. F. Daganzo, “Requiem for second-order fluid approximations of traffic flow,” *Transp. Res. Part B Methodol.*, vol. 29, no. 4, pp. 277–286, 1995.

[102] A. Aw and M. Rascle, “Resurrection of ‘Second Order’ Models of Traffic Flow,” *SIAM J. Appl. Math.*, vol. 60, no. 3, pp. 916–938, 2000.

[103] M. Papageorgiou, “Some remarks on macroscopic traffic flow modelling,” *Transp. Res. Part A Policy Pract.*, vol. 32, no. 5, pp. 323–329, 1998.

[104] D. E. Kirk, *Optimal Control Theory: An Introduction*, vol. 17, no. 3. 2004.

[105] C. F. Daganzo, “Urban gridlock: Macroscopic modeling and mitigation approaches,” *Transp. Res. Part B Methodol.*, vol. 41, no. 1, pp. 49–62, 2007.

[106] J. a. Laval and L. Leclercq, “Continuum Approximation for Congestion Dynamics Along Freeway Corridors,” *Transp. Sci.*, vol. 44, no. 1, pp. 87–97, 2009.

[107] C. Osorio and M. Bierlaire, “A simulation-based optimization framework for urban traffic control,” *Environ. Eng.*, pp. 1–29, 2010.

[108] R. Balakrishna, Y. Wen, M. Ben-Akiva, and C. Antoniou, “Simulation-Based

Framework for Transportation Network Management in Emergencies," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2041, pp. 80–88, 2008.

[109] T. Weise, "Global Optimization Algorithms–Theory and Application," *URL http//www. it-weise. de, Abrufdatum*, vol. 1, p. 820, 2009.

[110] M. Girianna and R. Benekohal, "Using Genetic Algorithms to Design Signal Coordination for Oversaturated Networks," *J. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 117–129, 2004.

[111] B. Park, C. Messer, and T. Urbanik, "Traffic Signal Optimization Program for Oversaturated Conditions: Genetic Algorithm Approach," *Transp. Res. Rec.*, vol. 1683, no. 1, pp. 133–142, 1999.

[112] B. R. Chilukuri, J. A. Laval, and A. Guin, "Optimal Ramp Metering with GA based Parameter Optimization," in *Transportation research Board*, 2015, vol. 4750.

[113] J. Grefenstette, "Optimization of Control Parameters for Genetic Algorithms," *IEEE Trans. Syst. Man. Cybern.*, vol. 16, no. 1, pp. 122–128, 1986.

[114] X. Y. X. Yang, L. C. L. Chu, and W. Recker, "GA-based parameter optimization for the ALINEA ramp metering control," in *Proceedings. The IEEE 5th International Conference on Intelligent Transportation Systems*, 2002, pp. 627–632.

[115] I. Wolfram Research, "Mathematica, Version 10.0.1." 2014.

[116] A. Gosavi, *Simulation-Based Optimization: Parametric Optimization Techniques and Reinforcement Learning*, vol. 1. 2003.

[117] I. Wolfram Research, "Functional Programming," 2014. [Online]. Available: https://reference.wolfram.com/language/guide/FunctionalProgramming.html.

[118] R. Storn and K. Price, "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces," *J. Glob. Optim.*, pp. 341–359, 1997.

[119] G. I. of Technology, "PACE: A partnership for an Advanced Computing Environment." [Online]. Available: http://www.pace.gatech.edu/. [Accessed: 11-

Dec-2015].

[120] W. L. Winston and M. Venkataramanan, *Introduction to Mathematical Programming: Operations Research*, vol. 32, no. 1. 2002.

[121] S. M. Robinson, *Springer Series in Operations Research and Financial Engineering*. 2006.

[122] K. Cunningham and L. Schrage, "LINGO: The modeling language and optimizer," *LINDO Syst. Chicago, Illinois*, p. 869, 2013.

[123] A. K. Ziliaskopoulos, "A Linear Programming Model for the Single Destination System Optimum Dynamic Traffic Assignment Problem," *Transportation Science*, vol. 34. pp. 37–49, 2000.

[124] C. Roncoli, M. Papageorgiou, and I. Papamichail, "Traffic flow optimisation in presence of vehicle automation and communication systems – Part I: A first-order multi-lane model for motorway traffic," *Transp. Res. Part C Emerg. Technol.*, vol. 57, pp. 241–259, 2015.

[125] S. Hoogendoorn, R. Hoogendoorn, M. Wang, and W. Daamen, "Modeling Driver, Driver Support, and Cooperative Systems with Dynamic Optimal Control," *Transp. Res. Rec.*, no. 2316, pp. 20–30, 2012.

[126] B. Cantwell and T. Moulden, "Introduction to Symmetry Analysis," *Appl. Mech. Rev.*, vol. 57, no. 1, p. B4, 2004.

[127] E. Buckingham, "On physically similar systems; Illustrations of the use of dimensional equations," *Phys. Rev.*, vol. 4, no. 4, pp. 345–376, 1914.

[128] M. Malinen, S. R. Duncan, T. Huttunen, and J. P. Kaipio, "Feedforward and feedback control of ultrasound surgery," *Appl. Numer. Math.*, vol. 56, no. 1, pp. 55–79, 2006.

[129] M. Desmurget and S. Grafton, "Feedback or feedforward control: End of a dichotomy," in *Taking action: Cognitive neuroscience perspective on intentional acts*, 2003, pp. 289–338.

[130] J. C. Walchko, J. S. Kim, K. W. Wang, and E. C. Smith, "Hybrid feedforward-

feedback control for active helicopter vibration suppression," in *Annual Forum Proceedings - AHS International*, 2007, vol. 2, pp. 1260–1276.

[131] R. M. Reddy, S. Member, I. M. S. Panahi, and S. Member, "Comparison between Hybrid Feedforward-Feedback, Feedforward, and Feedback Structures for Active Noise Control of fMRI Noise," *IEEE EMBS Conf.*, pp. 266–269, 2008.

[132] A. L. A. Luo, X. X. X. Xu, L. F. L. Fang, H. F. H. Fang, J. W. J. Wu, and C. W. C. Wu, "Feedback-Feedforward PI-Type Iterative Learning Control Strategy for Hybrid Active Power Filter With Injection Circuit," *IEEE Trans. Ind. Electron.*, vol. 57, no. 11, pp. 3767–3779, 2010.

[133] Wiki, "Pid Controller," 2014. [Online]. Available: http://en.wikipedia.org/wiki/PID_controller.

[134] D. Levinson, L. Zhang, S. Das, and a Sheikh, "Measuring the Equity and Efficiency of Ramp Meters. ," *Minnesota Dep. Transp.*, 2004.